

## **INFORMATION TO USERS**

**This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.**

**The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.**

**In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.**

**Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.**

**Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.**

# **U·M·I**

University Microfilms International  
A Bell & Howell Information Company  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
313/761-4700 800/521-0600



**Order Number 9419157**

**An examination of small class as a “gap reduction” strategy for  
achievement differences in groups of students, K-3**

**Bingham, Charles Steven, Ed.D.**

**The University of North Carolina at Greensboro, 1993**

**U·M·I**  
300 N. Zeeb Rd.  
Ann Arbor, MI 48106



AN EXAMINATION OF SMALL CLASS AS A "GAP REDUCTION" STRATEGY  
FOR ACHIEVEMENT DIFFERENCES IN GROUPS OF STUDENTS, K-3


by

Charles Steven Bingham

A Dissertation Submitted to  
the Faculty of the Graduate School at  
The University of North Carolina at Greensboro  
in Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Education

Greensboro  
1993

Approved by

  
\_\_\_\_\_  
Dissertation Advisor

BINGHAM, CHARLES STEVEN, Ed.D. An Examination of Small Class as a "Gap Reduction" Strategy for Achievement Differences in Groups of Students, K-3. (1993) Directed by Dr. C. M. Achilles. 123 pp.

The primary purpose of this research was to examine patterns of achievement among white and minority students in small (S), regular (R), and regular-with-aide (RA) class-size treatments over varying lengths of time and at varying start-up years, grades K, 1, 2, and 3. Re-analyzing norm- (NRT) and criterion-referenced (CRT) achievement test data from Tennessee's Student Teacher Achievement Ratio (STAR) Project, the researcher: (1) identified appropriate subsamples of students in each grade and class-size; (2) calculated the white and minority subsamples' mean scaled scores and standard deviations for each achievement measure by length of time in each class-size treatment at each grade; (3) tabulated the white and minority subsamples' mean scaled scores for each achievement measure by length of time in each class-size treatment at each grade level; (4) tabulated a white-minority achievement gap by grade by length of time in each class-size treatment; (5) tabulated achievement differences between minority (S) and white (R) subsamples by grade by length of time in class; (6) calculated effect sizes to measure the educational importance of selected mean score achievement differences between subsamples; and (7) charted histograms to compare the achievement of white (R) and minority (R) and (S) class-sizes.

The results for minority and white achievement for all years and class-sizes indicated that subsamples in (S) tended to obtain higher NRT and CRT mean scaled scores in reading and mathematics than either (R) or (RA) subsamples. Analyses consistently revealed a pattern of comparatively high 1 year mean scores with diminishing 2 years and 3 years scores followed by the greatest achievement in 4 years subsamples. The major difference between the minority and white pattern was the magnitude of the between-treatment subsample means: differences for white

subsamples tended to be compressed relative to differences between minority subsamples, that is, class-size appeared to make a bigger difference for minority students than for white students. Differences between minority and white achievement for all years and class-sizes tended to be smallest for the 2 years and largest for the 3 years time in treatment. Achievement differences between minority (S) classes and white (R) classes for all years revealed the smallest gaps at grade K and grade 1, 2 years. Compared to white-minority differences displayed within any one class-size, the gaps that resulted from white (R) and minority (S) were smaller on every achievement measure and at every grade level. Achievement differences between white (R) and minority (R) and (S) cohorts that matriculated synchronistically with the four years of STAR class-size treatment were smallest after one year (CRT outcomes) or two years (NRT outcomes).

The findings of this re-analysis supported the following conclusions: (1) Small class-size may be an effective white-minority achievement gap reduction strategy in the primary school years; (2) Benefits accrue initially in greater measure to minorities than to whites; (3) Optimally beneficial treatment begins no later than grade 1 (preferably in K); (4) Optimally beneficial treatment lasts at least two years; (5) The major benefit of small-classes for minorities appears to be in preventing rather than remediating achievement disadvantage; (6) The differential effect of small class-size for minorities appears to "fade" following two years treatment or in grade 3; and (7) As it applies uniquely to small class-size, the "fade" phenomenon of early treatments can be neither confirmed nor contradicted by this re-analysis.

APPROVAL PAGE

This dissertation has been Approved With Distinction by the following  
committee of the Faculty of the Graduate School at The University of  
North Carolina at Greensboro:

Dissertation Advisor C. M. Acquila

Committee Members Alab L. R. ...

[Signature]

[Signature]

11/9/93  
Date of Acceptance by Committee

11/2/93  
Date of Final Oral Examination



## ACKNOWLEDGMENTS

To Dr. C. M. Achilles, my committee chairman, I express my sincere appreciation for his support, encouragement and guidance. I also wish to thank my other committee members, Dr. Dale Brubaker, Dr. Roy Forbes, and Dr. David Strahan. For technical assistance in the computer analysis of STAR data I am indebted to Mr. DeWayne Fulton of the Tennessee Center for Excellence for Research in Basic Skills, Tennessee State University, Nashville, TN. Additionally, Dr. Barbara Nye, Center Director, and Dr. Jayne Zaharias, Research Assistant, assisted and encouraged this work. (Results expressed in this study do not necessarily reflect the position of the Center or its staff.) Thanks also to personnel in the Knox County, TN schools for confirming grade retention within a portion of the STAR data. I am also indebted to Mr. Patrick Harmon, Research Associate at the SouthEastern Regional Vision for Education (SERVE) for his expertise and advice in matters statistical. For contributing generally to my research interest, I thank my former mentors, Dr. Jack Taylor, Director of the Center for Music Research, Tallahassee, FL, and Dr. Clifford Madsen, Professor of the School of Music at the Florida State University. Finally, I am indebted to my wife, Alida, my son, Christopher, my daughter, Caitlin, my mother, Barbara, my brother, John, and my sister, Bobbi Lynn, for their moral support as I have endeavored to write this dissertation.

## TABLE OF CONTENTS

	Page
APPROVAL PAGE .....	i i
ACKNOWLEDGEMENTS .....	i i i
LIST OF TABLES .....	v i
LIST OF FIGURES .....	v i i
 CHAPTER	
I. INTRODUCTION.....	1
Statement of the Problem.....	4
Purpose of the Study .....	5
Research Question .....	6
Conceptual Base.....	6
Significance and Importance .....	1 6
Organization of the Rest of the Study .....	1 9
II. REVIEW OF RELATED LITERATURE.....	2 0
What Is the White-Minority Achievement Gap and Its History? .....	2 0
What Theories Explain the White-Minority Achievement Gap?.....	2 4
What Are Some Attempts to Reduce the White-Minority Achievement Gap?.....	3 3
Does Small Class-Size Reduce the White-Minority Achievement Gap?.....	6 7
III. METHODOLOGY.....	8 1
STAR Subjects .....	8 1
STAR Instruments and Procedures.....	8 2
Procedures Employed in the Present Re-Analysis.....	8 4
IV. RESULTS.....	8 6
Research Question (1).....	8 6
Research Question (2).....	8 7
Research Question (3).....	8 9
V. DISCUSSION.....	9 2
Summary of Results.....	9 2
Conclusions.....	9 5
Implications of the Results .....	9 8
Recommendations for Future Study .....	1 0 1
BIBLIOGRAPHY .....	1 0 4

**APPENDIX A. Achievement Measures Total n of Cases by Time in Treatment  
by Class-Size by Race.....1 17**

**APPENDIX B. Achievement Measures Standard Deviations by Time in  
Treatment by Class-Size by Race .....1 17**

**APPENDIX C. Tables of Additional Analyses of STAR Data Pertinent to  
Small Class as a "Gap Reduction" Strategy for Achievement Differences.....1 17**

## LIST OF TABLES

	Page
Table 1.....	33
Outline of Theories Explaining the White-Minority Achievement Gap	
Table 2 .....	86a
Minority Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Small Class-Size	
Table 3.....	87a
Minority Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular Class-Size	
Table 4.....	87b
Minority Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular/Aide Class-Size	
Table 5.....	88a
White Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Small Class-Size	
Table 6.....	88b
White Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular Class-Size	
Table 7.....	88c
White Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular/Aide Class-Size	
Table 8.....	89a
The White-Minority Achievement Gap: Differences in Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Small Class-Size	
Table 9.....	90a
The White-Minority Achievement Gap: Differences in Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular Class-Size	
Table 10 .....	90b
The White-Minority Achievement Gap: Differences in Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular/Aide Class-Size	
Table 11 .....	90c
Minority Small-Class and White Regular-Class Subsample Differences in Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Class	

LIST OF FIGURES

	Page
<b>Figure 1.....</b>	<b>91a</b>
<b>Comparison of SAT Reading Mean Scaled Score Scores for White (R) and Minority (R) and (S) Subsamples That Began in K and Matriculated All Four Years of STAR</b>	
<b>Figure 2.....</b>	<b>91b</b>
<b>Comparison of BSF Reading Mean Percent Passing for White (R) and Minority (R) and (S) Subsamples That Began in K and Matriculated All Four Years of STAR</b>	
<b>Figure 3.....</b>	<b>91c</b>
<b>Comparison of SAT Math Mean Scaled Scores for White (R) and Minority (R) and (S) That Began in K and Matriculated All Four Years of STAR</b>	
<b>Figure 4 .....</b>	<b>91d</b>
<b>Comparison of BSF Math Mean Percent Passing for White (R) and Minority (R) and (S) That Began in K and Matriculated All Four Years of STAR</b>	

## CHAPTER I

### INTRODUCTION

One of the greatest challenges facing U. S. education today is the differences in achievement among students of different ethnic groups. As a nation, we have ostensibly aimed to educate all of our youth; yet evidence exists which suggests that we have neither uniformly attempted nor succeeded in the effort. One need only visit East St. Louis (IL), Camden (NJ), or Halifax County (NC) to see just how different school can be for poor and minority-race children as opposed to middle-class and white children. Among students subjected to such inequities the real surprise is not that an achievement differential exists but that the differential is not more pronounced than it is! Prerequisite to redressing inequity, however, is the need to face a simple truth: "The notion here somehow is that we educate all kids the same. But somehow, Black kids, Brown kids, and poor kids don't learn as much. That is a serious misconception. In fact, we do not educate all children the same way (Haycock, 1990, p. 53)."

Supported by Edmonds' (1979a) and Lezotte's (1989) Effective Schools' research suggesting that some schools tend to educate students successfully without regard to students' categorical membership, enlightened educators now believe that schools can make a positive, measurable difference in student achievement regardless of ethnicity. Operationally defined, students within the top quartile of achievement would as likely be children of color and poverty as children who are white and economically advantaged (Sudlow, 1985). The vestige of educators' believing (and behaving) otherwise, one may argue, is the remaining "achievement gap" between white and minority students. And although at an historic nadir, the white-minority

achievement gap is still unacceptably large and troubling (Jennings, 1992; Levin, 1990; Mullis, Owens, & Phillips, 1990).

The size and nature of the achievement gap depend on a combination of the indicator, the location, and the minority group in question. Data compiled by the Sandia National Laboratory (1991) are illuminative: The Scholastic Aptitude Test (SAT) 1991 national mean score for Asian-Americans, for example, was actually higher than that of their white classmates. Similarly, retention and college attendance data indicated that there is, in fact, no achievement gap with respect to the Asian-American minority student. Black and Hispanic student data, on the other hand, demonstrated wide disparities. Using again the 1991 SAT as an indicator, the mean score for black students was approximately 200 points below the white student mean while Hispanic students experienced, on average, a 130 point disadvantage. Similarly, black and Hispanic minorities were shown to drop out of school at proportionately higher rates than their white classmates. The data indicated that nearly 80% of white students complete high school "on time," and about 88% do so by age 25. Alternately, only 70% of black students and 50% of Hispanics graduate "on time." By age 25, the graduation rate for blacks has climbed to about 82% (still 6% lower than whites), while only 60% of Hispanics have acquired their diplomas. When location is factored in, however, the data showed that the major gaps in white-minority achievement were occurring in the large urban school districts and could be as readily explained as a function of poverty as of ethnicity.

Other databases have shown equally severe educational achievement disadvantages of minority children. The Southern Regional Education Board (SREB, 1986) reported that minority students are often "two to three or more grade levels behind on achievement measures. On National Assessment Tests, the reading achievement levels of Black eleventh grade students is basically the same as for White

seventh graders (p. iv)." (Except in direct quotation of an author whose practice is otherwise, this writer will capitalize neither 'white' nor 'black' in this dissertation.) Results of the National Assessment of Educational Progress (NAEP) reading and mathematics tests for students grades 4, 8, and 12 (Mullis, Owens, & Phillips, 1990) demonstrated large white-minority differences in mean scores at every grade level and on both tests. A preliminary report of the 1992 NAEP assessment stated: "Although average performance for white students increased at each grade level . . . achievement for black and Hispanic students increased only at grade 12, and the gaps between whites' and blacks' performance remained substantial (Rothman, 1993, p. 23)." Differences abound, too, in terms of placement in special educational programs (Hathaway, 1990, p. 224): black students are approximately three times as likely to be in a class for the educable mentally retarded but only half as likely to be in a class for the gifted and talented; only about one-third of the estimated 2.7 million limited English speaking proficient (LEP) students aged 5 to 14 receive special help commensurate with their linguistic needs; black and Hispanic students are two to three times as likely to be suspended or expelled and only half as likely to be enrolled in courses that lead to college preparation. Consequently, black and Hispanic students continue to be significantly underrepresented in university degree programs (Marks, 1985; Sandia National Laboratories, 1991).

Black males appear to fare particularly poorly vis-a-vis their white classmates. Likened to an "endangered species" (Wright, 1992), black men, who make up just six percent of the U. S. population, are now three percent of college student enrollment and 47% of America's prison population (Hodgkinson, 1992, p. 3). Since 82% of America's prisoners, each of whom costs about \$22,500 annually, are high school dropouts (Hodgkinson, 1992, p. 3), one need not be a rocket scientist to arrive at the inescapable conclusion that if schools could but improve the achievement of the



black male subgroup alone, the savings to society may be profound. In summary, whether one considers scores on norm-referenced standardized tests (NRTs) of achievement, dropout data, placement in special educational programs, or post-secondary school status, minorities lag behind their white counterparts in educational achievement.

Educators must seek a reasonable option for reducing this disparity that is at once widely available, cost efficient, easily implemented and operational as a public school option.

#### Statement of the Problem

Although the size of the white-minority achievement gap has been reduced in recent years, the gap reduction has been small and the time for achieving it large. For example, it required 17 years to get a 15 point reduction in the NAEP reading scores for fourth grade students and 13 years to achieve an 11 point reduction in mathematic scores for twelfth grade students (Mullis et al., 1990). Moreover, there is evidence to suggest that without intervention, the achievement gap for any one student may increase dramatically over the course of the school career. Haycock (1990) shares one state's experience:

In California, at the first grade level we see very few, if any, differences in actual achievement between minority youngsters and other youngsters--generally no more than about 10 percentage points. By the time those youngsters reach sixth grade, the gap has grown quite a bit larger--many times on the order of about 30 percentage points. By the time those youngsters reach the twelfth grade, if they reach the twelfth grade at all, there are gaps of about 60 percentage points on tests like the Comprehensive Test of Basic Skills between the average Anglo youngster and the average Black or Hispanic youngster (p. 54).

Accordingly, education researchers are challenged to investigate early intervention strategies which promise to reduce the white-minority achievement gap. Among the fruits of such an investigation might be an enhanced understanding of all variables

involved, including the comparative effects of duration and incipience of treatment (x) on student achievement. Simply stated, policy must be guided by research that answers the question: If x, how much and when?

#### Purpose of the Study

Small class-size in the primary years of school, as suggested by Tennessee's Student Teacher Achievement Ratio (STAR) Project (Word, Johnston, Pate-Bain, Fulton, Zaharias, Achilles, Lintz, Folger, & Breda, 1990), may be an effective strategy for reducing the white-minority achievement gap. Drawing from the database generated by Project STAR, a Lasting Benefits Study (LBS) report (Nye, Achilles, Zaharias, Fulton, & Wallenhorst, 1992a) and a report of Project CHALLENGE (Nye et al., 1992b) tend to corroborate the original STAR findings with respect to closing the white-minority achievement gap. (A review of Project STAR and other pertinent studies will be accomplished in Chapter II, "Review of Related Literature.") Despite the minority student achievement benefits attributed to small class-size, however, the variables of (a) duration of (i.e., time spent in) "treatment" and (b) incipience of (i.e., year of first exposure to) "treatment" in a controlled experimental study remain unaccounted for. (Neither LBS nor Project CHALLENGE are experiments.) Essentially missing from the knowledge base, then, is a quantitative description of student achievement as a function of length of time in and school year(s) of exposure to class-size treatments. Accordingly, the primary purpose of this study was to examine patterns of achievement among white and minority students in small, regular, and regular-with-aide class-size treatments over varying lengths of time and at varying start-up years, grades K, 1, 2, and 3. The researcher was guided by STAR, LBS, and CHALLENGE findings (Achilles, Nye, Zaharias, Fulton, & Bingham, 1993) which suggest that (a) earlier intervention may be more effective than later intervention, (b) the benefits of intervention may be cumulative, and (c) small class-

size treatment may be more effective in preventing than in remediating the white-minority achievement gap .

#### Research Questions

(1) What patterns of minority student achievement (reading and mathematics Norm- Referenced Test (NRT) mean scaled scores and Criterion-Referenced Test (CRT) average percentage passing) are associated with exposure to small (S), regular (R), and regular-with-aide (RA) class-sizes for one year (grades K, 1, 2, or 3), two years (grades K+1, 1+2, or 2+3), three years (grades K+1+2 or 1+2+3) and four years (grades K+1+2+3)?

(2) What patterns of white student achievement (reading and mathematics NRT mean scaled scores and CRT average percentage passing) are associated with exposure to small (S), regular (R), and regular-with-aide (RA) class-sizes for one year (grades K, 1, 2, or 3), two years (grades K+1, 1+2, or 2+3), three years (grades K+1+2 or 1+2+3) and four years (grades K+1+2+3)?

(3) What patterns of differences in white-minority student achievement (reading and mathematics NRT mean scaled score differences and CRT average percentage passing differences) are associated with exposure to small (S), regular (R), and regular-with-aide (RA) class-sizes for one year (grades K, 1, 2, or 3), two years (grades K+1, 1+2, or 2+3), three years (grades K+1+2 or 1+2+3) and four years (grades K+1+2+3).

#### Conceptual Base

The conceptual foundation of this study builds on (1) theories advanced to explain the existence of the national white-minority (particularly black) achievement gap, (2) the Project STAR, LBS, and CHALLENGE research suggesting that small class size may be an effective intervention in reducing the gap, and (3) theories advanced to explain how reduced class size actually produces achievement gains for participating

students. The implication is that by knowing why the achievement gap exists, we can more effectively target interventions that will reduce it. Specifically, this section suggests that small classes provide a setting in which family and community effects, shown to be disproportionately pernicious among minority students, are offset and the relative impact of school programs and practices enhanced.

#### Why an Achievement Gap?

Some research supports comparative cultural differences as an explanation for why certain minorities lag in school achievement relative to majority whites. Ogbu (1992) explains that minorities may be classified as either (a) autonomous, (b) immigrant or voluntary, or (c) castelike or involuntary. His research suggests that involuntary minorities, people originally brought into the United States or any other society against their will, experience more difficulties in school learning and performance partly because of the relationship between their cultures and the mainstream culture (p. 9). In particular, a phenomenon called "cultural inversion" produces oppositional (with respect to the dominant white culture) behavior. Cultural inversion is described as the tendency for involuntary minorities to regard certain forms of behavior, events, symbols, and meanings as inappropriate for them because these are characteristic of white Americans (p. 8). For American Indians, blacks, and Puerto Ricans who consider themselves "a colonized people" achieving well in school is tantamount to losing their identity, to "acting white". Lemann (1986) also underscores the cultural difference theory when he argues that the distinctive culture of blacks is their greatest barrier to progress (p. 35). Lemann suggests that the increasing isolation of the black underclass is a primary source of the problem:

The plight of the ghetto is due to two mass migrations: The first was from the rural South to the urban North and numbered in the millions during the forties, fifties and sixties before ending in the early 1970's. This migration brought the Black class system to the North virtually intact, though the underclass

became more pronounced in the cities. The second migration began in the late sixties, a migration out of the ghettos by members of the Black working and middle classes, who had been freed from housing discrimination by the civil rights movement. Until then the strong leaders and institutions of the ghetto had promoted an ethic of assimilation (if not into White society, at least into a Black middle class) for the underclass, which worked up to a point. Suddenly most of the leaders and institutions (except criminal ones) left, and the preaching of assimilation by both Blacks and Whites stopped. What followed was a kind of free fall into what sociologists call social disorganization (p. 35).

Theories related to dyfunctionalities within black families are also advanced as explanatory of the achievement gap. McGhee (1985) sees three major problems facing black families today: (a) divorce and separation, (b) teenage pregnancy, and (c) nonemployment and unemployment. Affecting blacks in statistically disproportionate numbers, these problems tend to result in environments that are impoverished and nonsupportive of academic achievement. Staples (1987) suggests that the dominant force in the decline of the black nuclear family has its genesis in structural conditions. Despite evidence showing the willingness of black women to fulfill their normative role, data indicate that black men are thwarted in their efforts to be economic providers and family leaders. Prerequisite to providing, leading, and serving as a role model for youth is gainful employment, a condition which too frequently eludes the black male.

A third line of reasoning that suggests that blacks are genetically inferior (Jenson, 1969) or at least behave as though they believe that they are inferior (Howard & Hammond, 1985). Due to methodological flaws and reliance on the discredited work of Cyril Burt, Jenson's (1969) theory that African-Americans, as a population, score significantly lower on IQ tests than the white population as a consequence of their genetic heritage has been largely discounted. However, Howard and Hammond (1985) suggest that many African-Americans behave as if Jenson were right. Relating the inferiority argument to the power of expectations, they

hypothesize that black achievement problems are caused by a tendency to avoid intellectual competition resulting from their internalizing society's projection of black intellectual inferiority. A vicious spiral is thus created: Imputed genetic intellectual inferiority leads to data confirming poorer performance which intensifies the worry surrounding the issue which leads to poorer performance and diminished achievement. Alluding to the profound effect of expectancy on behavior and cognition, Howard and Hammond conclude:

Expectations of Black inferiority are communicated, consciously or unconsciously, by many Whites, including teachers, managers, and those responsible for the often demeaning representations of Blacks in the media. These expectations have sad consequences for many Blacks, and those whose actions lead to such consequences may be held accountable for them. If the people who shape policy in the United States, from the White House to the local elementary school, do not address the problems of performance and development of Blacks and other minorities all Americans will face the consequences: instability, disharmony and a national loss of the potential productivity of more than a quarter of the population (p. 21).

The relationship between race and expectations is also demonstrated in Jaeger's (1982) research on a judgmental method for establishing standards on North Carolina high school competency tests. In exploring the relationships between background characteristics (e.g., race, gender, age, education, etc.) of judges and their recommended standards, Jaeger reports that race is the only demographic variable that is significantly predictive of judges' recommended standards:

Black judges had a mean recommended standard on the reading test that was almost 12 points lower than the mean recommended by white judges. The difference is more than a standard deviation. The pattern is similar for those rating items on the Mathematics test. The mean standard recommended by black judges was eight points lower than the mean standard recommended by white judges; this difference is more than half a standard deviation (p. 471).

Left to speculate as to the cause of the wide disparity between the recommended standards of white and black respondents, one is forced to consider the possibility that

blacks simply held lower expectations of what is attainable. That such expectations may reflect the black experience in America further supports Howard and Hammond's (1985) hypothesis.

Sizemore (1990) epitomizes the critical theory perspective, that the white-minority achievement gap can be explained in terms a social power differential and as a nation, we have simply lacked the political will to address the needs of minority students. She asserts:

We have the technology to rescue at-risk learners from failure. It is no longer an educational question. It is a political problem. Politics is the management of the conflict of groups at war over scarce resources. Power is the central concept. When those who have power decide that we should educate the at-risk students, we will do so on a grand scale. Maybe that time has come. For if the baby boomers do not educate the young minority students in school now, there may be no one to pay for their old age pensions (p. 51).

In his compelling expose' of inequities within the U. S. educational system, Kozol (1991) echos Sizemore and extends her argument. During a recent interview in which his book, Savage Inequalities, is discussed (Scherer, 1993), Kozol states:

The problem is not that we don't know what works. The problem is that we are not willing to pay the bill to provide the things that work for the poorest children in America. . . I chose that title [Savage Inequalities] because I was sick of powerful people suggesting that there was some kind of essential savagery in poor black children in America. I wanted to make clear that if there is something savage in America, it is in the powerful people who are willing to tolerate these injustices (pp. 8-9).

In accepting the achievement gap between mainstream students and minority students, Hilliard says that Americans have "rationalized" and tend to blame the gap on false but popular causes such as socioeconomic status and cultural diversity. "The real cause of the achievement gap is the differential treatment that students receive. . . We always talk about the achievement gap, not the treatment gap (Willis, 1993)." Sizemore,

Kozol, and Hilliard theorize that poverty, powerlessness, and minority status often go hand in hand and that the national achievement gap is the result of a power imbalance. However, when those in power decide that the gap is intolerable, then ways to reduce the gap will be found. That is, cultural differences will be minimized, dysfunctional families made functional, and expectations for minority children raised. One possible means is the focus of the second part of this section.

#### Small Class-Size as an Early Intervention

Results of the Student/Teacher Achievement Ratio (STAR) Project and its progeny, the Lasting Benefits Study (LBS) and Project CHALLENGE, suggest that small class size in the early primary grades (K-3) may be effective in closing the white-minority achievement gap. (Both STAR, LBS, and CHALLENGE will be examined in detail in Chapter II, "Review of Related Literature.") Funded by the Tennessee legislature, STAR was a large-scale experiment in which kindergarten students and teachers were randomly assigned to small and large classes within each participating school. Students remained in these class-types for four years. At the end of each year, they were measured in reading and mathematics by standardized norm and curriculum-based tests. The results were definitive: (a) a significant benefit accrued to students in reduced-size classes in both subject areas and (b) there was evidence that minority students in particular benefited from the smaller class environment, especially when curriculum-based tests are used as the learning criteria. Longitudinal analysis indicated that students in small classes outperformed their peers in kindergarten classes of regular size and also gained more in reading outcomes the second and subsequent years (Finn & Achilles, 1990). Summarizing the findings pertinent to the present study, Finn and Achilles assert:

In addition to an overall class-size effect, there is strong indication that the performance of minority students is enhanced in the small-class setting. This



important outcome is statistically confirmed only in inner-city and suburban areas, but the same trend is seen in urban and rural schools as well. Also, minority students in the longitudinal subsample experienced greater relative growth than white students in the second year of small-class participation (p. 574).

The LBS is the longitudinal follow-up study of Project STAR conducted to determine if there are continuing achievement benefits for STAR students after they return to regular-size classes. Analysis of fourth grade data showed clear and consistent results from norm- and criterion- referenced tests. Students previously in a small-size STAR class demonstrated statistically significant advantages over students in large-size classes. Nye et al. (1992a) report: Although Project STAR results on K-3 test scores indicated that the most significant S [small] class advantages were found with minority students, the LBS fourth-grade results showed that the positive effects of small classes were equally significant for minority and white students, in all school locations (p. 9).

The administration of the tests in the fourth-grade excluded a large group of students, a high percentage of which were urban minority due to a testing change. The group was returned to the analysis in grades 5 and later. (The minority sample diminished by nearly a third.) Given this fact as well as the assertion by the researchers that "The class-size intervention effect may be expected to diminish in some subsequent school years and to reappear in others (p. 10)," the need for continued longitudinal analysis using the Project STAR database was suggested. Accordingly, Nye et al. (1992b) analyzed test data from fifth-grade students. Again, the data showed a statistically significant advantage for students who had been in small classes during the original STAR experiment two years after returning to regular-size classes. The researchers speculate:

Since LBS data show continuing benefits in pupil achievement after small-class involvement, will small-class involvement for one of two years (rather than STAR's four years) provide a sound base to help pupils get started well in school? If so, STAR results were strongest in K and 1, suggesting that these should, at minimum, be the years of the small-class intervention. The early primary heterogeneous classes provided by the STAR random assignment and STAR's seeming ability to help minority pupils close the achievement gap are promising areas for LBS analyses (p. 13).

A critical question, however, remains: What happens in small classes that would cause attending students, minority or otherwise, to achieve more?

#### How Does Small Class-Size Explain Achievement Gains?

Explanations for how small class-size contributes to achievement beyond what may be expected from regular-size classes have been posited. Finn and Achilles (1990, p. 575) suggest that three dimensions of school processes may be examined: First, enhanced teacher satisfaction due to reduced class-size may positively influence student motivation for learning. This theory is supported by the Glass, Cahen, Smith, and Filby (1982) and Shapson, Wright, Eason, and Fitzgerald (1980) research findings. Second, reduced class size may directly impact teacher-student interactions and optimize individual attention to any one student. Third, smaller class size may increase pupils' attention to and involvement with learning activities. Alternately, Finn, Fulton, Zaharias, and Nye (1989/1992) suggest that "reducing the number of students in a class does not impact performance directly but instead affects the processes that mediate achievement (p. 83, authors' emphasis)," while Nye et al. (1992b) assert that "small classes allow for more developmentally appropriate curriculum, instruction and parent involvement (p. 14)."

Mitchell, Beach, and Badarak (1989/1992) argue that STAR, like most class size research, suffers from its reliance on "'brute force empiricism'--testing whether there are differences in the achievement levels of large and small classes

without any clear theory of how class size reduction might be turned into achievement test scores gains (p. 36)." Although certain criticisms by Mitchell, Beach, and Badarak of STAR have been challenged (Achilles, Nye, Zaharias, Fulton, & Wallenhorst, 1992), they may be credited with identifying six explicit models/theories of how class size may impact student test performance. Three of these models assert a direct link between larger numbers of students in a classroom and declining achievement test scores. Arguing that adding more students to a class erodes the teacher's instructional effectiveness, the theories are (a) greater Instructional Overhead, (b) increased Student Interaction Time and (c) decreased access to Fixed Instructional Resources. A second group of three class size theories posit an indirect link between more students and altered classroom performance. Arguing that it is not the number of students, but factors associated with student assignment to large and small classes to which achievement effects may be attributed, the second group of theories are (d) Class Heterogeneity, (e) Instructional Pacing and (f) Student Grouping or Achievement Modeling. To determine the extent to which the STAR data confirmed or failed to confirm the six models, the researchers used regression analyses. They found in part that:

The Project STAR data strongly support the conclusion that the impact of class size on student achievement is the indirect result of differences in the pattern of student achievement in each classroom, rather than a direct effect of the number of children in each classroom. The most powerful model is the one which attributes classroom achievement to the independent effects of the least able and most able students--mediated either through teacher grouping practices or through the contagious effects of the academic aspiration of high and low students on their peers. This model makes class size appear to be much more important than it is. For some reason, low performing students are more often found in larger classes while their high performing counterparts are about equally distributed between large and small class settings, reducing the achievement level of regular classes while raising that in the smaller ones (authors' emphases, pp. 65,66).

Mitchell et al.'s conclusions are as confusing as they are compelling. In citing the primacy of indirect effects of class size on achievement, there is apparent agreement with Project STAR researchers. However, they also imply that STAR students were grouped into classes on the basis of ability. This is absolutely false. The very strength of Project STAR is the randomization of both teachers' and students' assignments into small- and regular-sized classrooms. Moreover, in concluding that "reducing the achievement level of regular classes while raising that in smaller ones" resulted from "low performing students [being] found in larger classes," Mitchell et al. use the results of small-class intervention (the outcome) to question the treatment. The logic of this seems questionable. Nevertheless, re-tests of the indirect models separating the sample into regular and small classes yielded a finding that, while only "interesting" to Mitchell et al., is extremely significant to this study:

It is interesting to note that the school intake factors--socioeconomic status, ethnicity, gender, and school type--are typically more powerful predictors of overall class achievement in the larger classes. This would indicate that teachers in small classes are better able to offset the effects of these family and community based factors and strengthen the relative impact of schooling programs and practices (p. 66).

Perhaps offsetting environmental effects when they are negative (e.g., discouragement from "acting white" by minority peer groups, social disorganization, dysfunctional families, imputed intellectual inferiority, low expectations, poverty and powerlessness) is exactly why small classes may give minority students "a leg up" relative to white students.

In summary, both STAR/LBS researchers and Mitchell et al. suggest that achievement gains for students in small classes may be due either to direct or indirect causes. Although Mitchell et al. have formulated some compelling theories with one, the Student Grouping/Modeling indirect model, being particularly powerful, the

implications of their findings do not fully align with the facts of the STAR experimental design, namely its randomization. The problem is, as Achilles et al. (1992) point out, "Their [Mitchell et al.] discussion of 'non-random' is based on the testing results at K. That is, they use the effects (what the study showed) to try to explain non-randomness (p. 3)." Nevertheless, Mitchell et al. may be credited with demonstrating why minorities, in particular, may benefit from small class size. The exact mechanism by which these benefits accrue remains a fertile field of inquiry.

#### Significance and Importance

Today, nearly 40 years after Brown vs. Topeka, and almost 30 years after the Elementary and Secondary Education Act of 1965 pledged our national commitment to equity and excellence in education, there remains a white-minority achievement gap. Desegregation, compensatory education, multicultural education, and most recently the effective schools movement, have been the institutional answers to those who seek to reduce the achievement gap. If one equates minority status with "at-riskness," (and in many locales, particularly urban areas, the comparison is valid), many other school interventions have been tried (Slavin, Karweit, & Wasik, 1993). In a larger sense, programs and services for at-risk children and their families that extend beyond the school (e.g., prenatal care, child care, health education, and substance-abuse prevention) have also been successful in mitigating the differences between the culturally advantaged and disadvantaged (Hamburg, 1992). Indeed, many indicators suggest that the gap is closing, but one may ask: Is it closing fast enough?

Clearly, if demographic trends and projections are any indication, the signposts say that our failure to attain white-minority achievement parity now will cost us dearly down the road. In terms of predicted changes in ethnic distribution and the effects on education and the workplace, the implications for continuing inequity are disastrous. Hodgkinson (1991, 1992) informs us that during the 1980 to 1990

period, the American population grew by 10%, an addition of 29 million people. Of that 29 million, 14 million were minorities. While the white population increased by 15 million, their share of the total U. S. population declined from 86% to 84%. Projecting changes in the youth population is even more provocative. As the total youth cohort moves from its current 64 million to 65 million in 2000, then down to 62 million, the minority component will increase dramatically, from 30% in 1990 to 38% in 2010. However, four states (New York, Texas, California and Florida) will have about one-third of the nation's youth, and more than half will be "minority". The real minority then becomes the non-Hispanic white youth. As Hodgkinson wryly queries, "What will we call 'minorities' when they are more than half of the population (1992, p. 6)?" Indeed, three years ago, 26 California cities contained no single majority racial or ethnic group.

It is the climate of increasing diversity that poses the biggest danger to education. Orum (1990) clearly articulates the problem: "The schools are being called upon to serve a growing proportion of ethnic- and language-minority-group children--the very children they have historically served least well (p. 83)." To the extent that teachers themselves, by their own ethnicity, contribute to reducing the white-minority achievement, the outlook is not good. Based on who is currently preparing to become teachers in schools of education, only 19% of the 2010 teaching staff will be nonwhite in California, while the numbers for Florida and Texas will be 15% and 22%, respectively. Bearing in mind that more than half the students in these three states will be nonwhite, the question is: Who will assist new white teachers in discovering what their diverse students are like (Hodgkinson, 1992, p. 9)?" The education of inner city youth also promises to become increasingly problematic as the flight to the suburbs continues, leaving huge pockets of minorities to deal with crime,

drugs, poverty, youth violence, and family disruption. Neither teachers nor administrators will be easily convinced to work in the inner cities of our nation.

Finally, projections about future jobs and the role of minorities argue for closing the achievement gap with urgency. Hodgkinson (1992, p. 9-10) says that the new workers in 2000 will be 85% combinations of immigrants, women, and minorities; only 15% will be white males. The kind of work that will be available is basically minimum wage occupations that can be performed by high school dropouts, such as janitor, clerk, fast-food worker and hotel-room maid, and well-paid occupations in technical or administrative positions (particularly in the health field) that require a college degree for entry. Given that minorities are such a large percentage of the new workers in the next 20 years, Hodgkinson predicts one of three scenarios: (1) If more minorities graduate from high school and take some community college course training, 1.2 million health technology jobs await them; (2) If minorities drop out of high school, 4.2 million new service jobs will be waiting, of which 3 million are minimum wage and can be done by high school dropouts. (Minorities are now over-represented in this job category.); (3) If minorities go to college and graduate, 3.5 million new professional jobs (lawyers, doctors, teachers, accountants, etc.) await them. (Minorities are now under-represented in these fields.) Clearly, the nation stands to benefit most if minorities choose options one and three. Without increased attention to closing the white-minority achievement gap, however, the likelihood of minorities being free to select the kind of work they will do is diminished.

Within the next 20 years, Americans will recognize that as the number of children declines as a percent of the U. S. population and people of color (non-Anglos) become the majority, we literally cannot afford to allow the achievement of the "new majority" youngsters to lag behind their white counterparts. We cannot wait 20 years

to do something about it. Preliminary research suggests that small class-size in the primary years of school may be an effective strategy for reducing the white-minority achievement gap. Obtained through the re-analysis of the STAR database, the findings of this study may contribute to education researchers' understanding of the interaction of ethnicity and the duration and incipience of class-size "treatment" with student achievement. More importantly, the findings may suggest to education policymakers guidelines for the more effective implementation of small class-size as a cost efficient public-school option.

#### Organization for the Rest of the Study

Following this chapter will be: (1) Chapter II, a review of literature related to the white-minority achievement gap focusing on (a) a definition of the white-minority achievement gap and its historical context, (b) theories purporting to explain the gap, (c) attempts to reduce the gap, and (d) small class-size as a gap reduction strategy; (2) Chapter III, the method used in the present study including (a) a description of STAR subjects, (b) STAR instruments and procedures, and (c) procedures used in this re-analysis; (3) Chapter IV, a report of the results of the present re-analysis; and (4) Chapter V, a discussion including (a) a summary of results, (b) conclusions, and (c) recommendations for future study.



## CHAPTER II

### REVIEW OF RELATED LITERATURE

Four questions are fundamental to a longitudinal study of white-minority achievement gap reduction attributed to small class-size: (1) What is the white-minority achievement gap and its history?; (2) What theories explain the white-minority achievement gap?; (3) What are some attempts to reduce the white-minority achievement gap?; and (4) Does small class-size reduce the white-minority achievement gap? Accordingly, the literature reviewed in this chapter provides a basis for addressing these four questions. The definition and historical context of the white-minority achievement gap are explored in the first section. The second section describes theories that might explain why minorities have historically lagged behind their White peers in scholastic achievement. Using theories outlined in the second section as organizers, the third section provides evidence for the white-minority achievement gap and illustrates a number of programmatic attempts to reduce it. Studies which suggest that the gap is reduced by small-class intervention in the primary years of school comprise the fourth section.

#### What is the White-Minority Achievement Gap and its History?

Education writers reviewed in this chapter seem to use and understand the phrase, "white-minority achievement gap" in remarkably similar ways. In a typical example, Martinez (1985) defined the white-minority achievement gap as "the ongoing disparity between the educational achievement of White and Asian students and the achievement of Blacks, Hispanics, American Indians, and students from other minority groups (p. 1)." Haycock and Brown (1988), Mullis et al. (1990), and Jennings

(1992) provide similar definitions. Martinez explained that, in discussions of the white-minority achievement gap, a distinction is commonly made between minority groups and underrepresented minority groups who lag behind whites and Asians in educational achievement and who are underrepresented educational achievers (author's emphasis). Blacks, Hispanics, and American Indians were seen to belong to the underrepresented group (Martinez, 1985, p. 1). In terms of measures of achievement, norm-referenced test (NRT) student scores (e.g., NAEP test scores) were typically utilized; criterion-referenced test (CRT) student scores, school attendance and drop-out data, retention and failure rates, and post-secondary status were employed less often. (Literature in a subsequent section of this chapter will illustrate these various criteria.) However, as the literature and research in this section demonstrate, the white-minority achievement gap cannot be fully understood or appreciated apart from its historical context.

In The Troubled Crusade: American Education 1945-1980, Ravitch (1983) argued persuasively that the white-black achievement gap is, in part, a result of resource inequities promulgated by schools' application of the 1896 Plessy v. Ferguson decision which upheld the "separate but equal" doctrine. In the segregated schools of South Carolina, for example, Ravitch reported that the number of white and black students was nearly equal in 1945; however, the state spent nearly three times as much for each white pupil as for each black pupil; the value of white property was six times the value of black school property; and the state spent 1/100 as much transporting black pupils to school as it did transporting white pupils. Moreover, federal funds allocated to the states for vocational education and for teacher training were apportioned inequitably between white and black schools: blacks, who were 21.4% of the pupil population in states with segregated systems, received only 9.8% of federal

dollars in the mid-1930s. Add to this injury the insult of White schools' routinely "handing down" used textbooks and equipment to the black schools (p. 121).

Ravitch documented a series of educationally significant judicial and congressional actions, including the 1954 Brown and 1955 Brown II decisions, the passage of the Civil Rights Act of 1964, and the Elementary and Secondary Education Act (ESEA) of 1965, aimed at redressing the inequities perpetrated by a majority white system on minority blacks. With Ornstein (1982), Ravitch also cited the effectiveness of the popular books Slums and Suburbs (Conant, 1961), The Culturally Deprived Child (Riessman, 1962), and The Other America (Harrington, 1963), in raising America's collective conscience regarding the plight of poor and minority youth. It remained, however, for the massive Equality of Educational Opportunity report, published in 1966 and popularly known as the Coleman Report, to reveal the cumulative effect of inequality on student achievement.

The Coleman Report (Coleman et al., 1966) was commissioned by the Civil Rights Act of 1964 as a survey concerning the lack of availability of equal educational opportunities for individuals by reason of race, color, religion, or national origin in public institutions at all levels in the United States. Led by James S. Coleman of Johns Hopkins University and Ernest Q. Campbell of Vanderbilt University, researchers surveyed some 4,000 public schools in the fall of 1965. They examined not only school inputs such as facilities, materials, curricula, and laboratories, but also analyzed outcomes. For the first time, the relationship of school resource allocation with student achievement, as measured by NRT scores, was addressed. Two of the major findings from the Coleman Report were: (1) The academic achievement of children from minority groups was one to two years behind that of whites at first grade; by twelfth grade, minority children were as much as three to five years behind their white peers;

and (2) Achievement seemed to be related to the student's family background rather than to the quality of the school. Ravitch suggested that although the second finding generated the most controversy, implying as it did the relative unimportance of schooling on student achievement, the first finding may be seen as the seed from which subsequent research on the white-minority achievement gap grew (p. 169).

Ravitch (1983) also argued that the findings of the Coleman Report and the 1967 U. S. Commission on Civil Rights' report, Racial Isolation in the Public Schools, contributed to a growing consensus that racial isolation was responsible for the white-black achievement gap. To close the gap meant that schools must comply with the intent of Brown and desegregate. In 1971, the Swann v. Charlotte-Mecklenburg decision was said to demonstrate the Supreme Court's resolve to use racial redistribution (busing) as a remedy for previously unconstitutional segregation. Ravitch reported, however, that social scientists soon began to challenge the assumptions of the integration policy. David J. Armor (1972) reviewed a number of studies and concluded that racial balance had not led to higher achievement or self-esteem among black students and had not improved race relations. Nancy St. John (1975) analyzed over 100 studies to determine the effects of school desegregation on children finding that there was no clear evidence that changes in the racial composition of the school closed or reduced the achievement gap between the races. Coleman (1975) himself released a study that concluded that court-mandated school desegregation contributed to the flight of whites from big cities and was thus a self-defeating policy. Ravitch reported that dissension on issues of race and education included blacks themselves. In Atlanta and Dallas, for example, the black community persuaded federal district judges not to impose system wide racial balancing and to seek instead improvement of their neighborhood schools. Ravitch concluded:

That blacks had begun to disagree openly with those who spoke on their behalf, that blacks did not feel that an institution was stigmatized by their participation in it, that black clients had found their own voice were indications not of a dissolution of the movement for black equality but of the success of the social revolution initiated by the Brown decision. In the quarter-century after Brown, race relations had changed dramatically. With the protection of court decisions and civil rights laws, blacks entered every walk of life, moved into formerly lily-white occupations, went to college in growing numbers, took an active part in political life, and destroyed forever the subservient, inferior role imposed by white racism in the past (p. 181).

By providing a historical context within which to examine the white-minority achievement gap, Ravitch's (1983) contribution is significant. She marshalls evidence to suggest that the long-standing inequities of resource allocation prevalent in U. S. schools prior to Brown, and still common in many districts today, may contribute to differential achievement outcomes for black and white children. (This suggestion forms the basis of a formal theory examined in the next section of this chapter.) However, to the extent that the social gains that have been achieved by blacks may be attributed to policy which, in turn, was influenced by research (e.g., Coleman et al, 1966), Ravitch's chronology represents for the research community a desirable phenomenon. Beyond the history of the white-minority achievement gap as a research interest, what theories have been advanced to explain its existence and persistence? This question is addressed in the next section.

#### What Theories Explain the White-Minority Achievement Gap?

Despite years of political and judicial activism in equity issues surrounding public education (some of which was reviewed in the previous section), education writers have been strangely silent in theorizing about the white-minority achievement gap. Although important education policies and programs (e.g., desegregation, multicultural education, Effective Schools) have been generated by the notion that "something" must be retarding the school success of minority students, the "somethings" have been systematically and comprehensively illuminated in only two

pieces of literature, Lansa and Potter (1984) and Sizemore (1990). Although self-referencing and ahistorical in tone, a third model (Haycock & Navarro, 1988) chronologically and conceptually bridges the Lansa and Potter and Sizemore frameworks.

At the 1984 National Conference on Desegregation in Postsecondary Education, Phillip J. Lansa and Jane Hopkins Potter presented a paper entitled, Building a Bridge to Span the Minority-Majority Achievement Gap. Lansa and Potter discerned six categories of theory that have been advanced to explain why black students lag behind whites in school achievement: (1) Environmental theory, (2) Cultural Deprivation theory (sometimes called 'hypothesis' and a subcategory of Environmental theory), (3) Coleman's Family Background theory, (also a subcategory of Environmental theory), (4) Jenson's Genetic theory, (5) Cultural-difference theory (sometimes called 'hypothesis'), and (6) Effective Schools theory (also a subcategory of Environmental theory).

According to Lansa and Potter, Environmental theory emerged in the 1950s and 1960s. Environmentalists were said to be divided at that time into two camps: (a) those who believed that the school determined achievement, and (b) those who felt that home and family were responsible. In the first camp were those educators and social scientists who sought to eradicate segregation and whose efforts resulted in the Brown decision of 1954. These individuals were said to adhere to a "segregation hypothesis." It held that schools could and should provide educational equality by creating environments in which students from diverse groups were free to interact and learn together (p. 2).

With the 1960s came the subscribers to the Cultural Deprivation theory. According to Lansa and Potter, underlying Cultural Deprivation is Piaget's theory that a child who, at an early age, is deprived of appropriate environmental stimuli lacks the

experiences needed for the development of intellectual skills and abilities (p. 4). Compensatory education programs, such as those created through the Elementary and Secondary Education Act of 1965, were said to have been engendered by Cultural Deprivation theory.

Lansa and Potter held that the work of James Coleman et al. (1966) countered those who saw schools as primarily responsible for students' achievement. According to Coleman, the single most important variable in determining a child's academic achievement was the educational and social background of the child's family (p. 6). In a re-analysis, Christopher Jencks et al. (1972) corroborated Coleman's original findings. The Coleman theory thus argued that compensatory education programs and busing as solutions to narrowing the white-minority achievement gap were ineffective.

A Genetic theory advanced by Arthur Jensen (1969) was also presented by Lansa and Potter as explanatory of the white-minority achievement gap. Jensen argued that genetic factors were strongly implicated in the average white-black intelligence differential. Like that of Coleman and Jencks, Jensen's work is said to minimize the impact of schooling, and particularly compensatory programs, as a means of ameliorating achievement differences between white and minority children (p. 6).

Lansa and Potter stated that Cultural-difference theory emerged in the late 1960s and early 1970s. They claimed that its advocates rejected the views of the geneticists and cultural deprivationists and argued that the culture of the school (school climate) must be reformed to make it more congruent with the diverse cultures of minority students. Neither completely attributing achievement to the home nor the school, the Cultural-difference proponents argued that interaction of the two is the significant factor (p. 7). Multiethnic and later, multicultural, education were seen as the programmatic outcomes of the Cultural-difference theory.

Finally, Lansa and Potter proposed that Edmonds' (1979) and Lezotte's (1989) Effective Schools research be seen as the embodiment of another theory to explain differences in white-minority achievement. Edmonds and Lezotte found that some schools were more successful than others in promoting the achievement of all students regardless of categorical membership. Neither genetics nor family background explained all of the achievement differences in schools whose characteristics included (a) leadership which focuses on the identification and diagnosis of instructional problems, (b) purposes which are clearly defined and recognized, (c) a safe and orderly environment, (d) high expectations for all students, and (e) regular monitoring by standardized tests (p. 11).

Lansa and Potter may be criticized on two counts: (1) Their review of literature was extremely scanty (only 13 references) and (2) they arbitrarily (and wrongly) interchanged the terms 'theory' and 'hypothesis' illustrating a possible flawed concept of what constitutes a theory. Nonetheless, Lansa and Potter may be credited not only with presenting a coherent historical overview of the issues and concerns pertinent to the white-minority achievement gap but with how policy and programs have been impacted by research.

In a report from the Achievement Council of California, Unfinished Business: Fulfilling Our Children's Promise, Haycock and Navarro (1988) provided another theoretical perspective. Their framework was two-part: (1) the schooling experience and (2) families and communities. Relative to the schooling experience, Haycock and Navarro asserted:

Into the education of poor and minority children, we put less of everything we believe makes a difference. Less experienced and well-trained teachers. Less instructional time. Less rich and well-balanced curricula. Less well-equipped facilities. And less of what may be most important of all: a belief that these youngsters can really learn (p. 18).



Haycock and Navarro specified five ways that the schooling experience of minorities systematically places them at disadvantage: (1) In elementary school, minorities are exposed to less demanding books and assignments and the lowest level conceptual skills. Worksheets are the instructional media of choice. In junior and senior high school, ability grouping and tracking cause minorities to fall further behind. (2) Those who teach in predominantly minority schools are more likely to have less experience and less education than those who teach in the suburbs. They are more likely to hold emergency credentials, teach out of field, and be hired as a long- or short-term substitute. They are not very likely to be minority. In Los Angeles, for example, only 17% of the teachers are minority compared to 49% of the students. Principals in minority schools also report being unprepared and are less credentialed than non-minority school principals. (3) Counselors in minority schools are overworked. Due to budget cutbacks in urban (minority) schools, counselors who used to have caseloads of 250 students are now charged with 450 to 750 students. (4) By their actions, teachers often manifest low expectations of minority students. Such actions include (a) assigning students to low reading groups, (b) using controlling, as opposed to instructional, questioning, (c) excusing late papers or missed assignments, (d) inflating grades, and (e) not assigning homework. (5) The educational facilities of minority and poor youngsters tend to be overcrowded, poorly-equipped, and more run-down than those of their non-minority peers.

Haycock and Navarro also argued that the families and communities of minority youngsters helped explain the achievement gap in at least five ways: (1) Low parental aspirations and level of education is agentic. In 1987, for example, 50% of white California parents held a Bachelor's degree, compared to 35% of black parents and only 14% of Latino parents. (2) Differential economic circumstances are causative. Minority parents tend to be poorer than white parents and consequently, less able to

provide enriching educational experiences beyond the offerings of the schools. (3) Distressed conditions within neighborhoods result in few positive role models for minority youth. Middle class blacks and Latinos have abandoned the cities for the suburbs, leaving in their wake high unemployment, empty churches, and disbanded civic organizations. A pervasive sense of hopelessness and fear contributes to the relatively low priority of education. (4) The mobility rate of minority families is high, that is, they tend to move around a lot, as much as two or three times in a year. There is not enough affordable housing causing people to move from project to project, relative to relative, often just "one step ahead of the bill collector." The affected children often become disruptive in the classroom and feel unconnected to the larger school community. (5) Minority peer influences minimize the importance of school. To achieve well in school is tantamount to "acting white" and therefore, to being disloyal to the minority peer group.

Haycock and Navarro's (1988) framework is at once more parsimonious and more expansive than the more complex model of Lanza and Potter (1984). Dispensing with theory imputing genetic differences as causative of the white-minority achievement gap, Haycock and Navarro suggested that only environmental factors are agentic. Thus focused, they drew from empirical and experiential sources to broaden the argument that the combination of disadvantaged homes and families and inequitable school conditions predispose minority youth to underachievement. Significantly, Haycock and Navarro's model foreshadowed Ogbu's (1992) research that suggested that involuntary minorities experience more difficulties in school learning and performance partly because of the relationship between their cultures and the mainstream culture. Consequently, Haycock and Navarro's framework may be viewed as a bridge to more contemporary critical perspectives.

In an article entitled, "Effective Education for Underachieving African-Americans," published as a chapter in Bain and Herman's (1990), Making Schools Work for Underachieving Minority Students: Next Steps for Research, Policy, and Practice, Barbara Sizemore stated that five categories of theory have been advanced in the literature to account for the underachievement of black students: (1) African-Americans are genetically inferior in intelligence; (2) African-Americans are culturally deprived; (3) African-Americans have deficient, indifferent, unstimulating, and immoral families, homes, and community environments; (4) African-American students are served by schools and/or school systems that are inefficient, underfunded, and ineffective; and (5) the larger social order dictates through its value systems a racial/caste system that perpetuates itself through the schools (Sizemore, 1985; 1989). Referencing Effective Schools literature, Sizemore suggested the first three categories of theory functioned to undermine the belief that all children can learn given sufficiently high expectations to do so (p. 41).

The first category was attributed primarily to the work of Arthur R. Jensen of Harvard University (Jensen, 1969, 1981; Sizemore, 1989). Jensen theorized that generally lower IQ test scores of blacks, as compared to whites, was genetically-based. Critics of Jensen asserted that his scientific methods were flawed and that he relied on the discredited work of Cyril Burt (Gould, 1981). Others suggested that science could not begin to differentiate the effects of heredity from environment until social conditions had been equal for both races for several generations (Bodmer & Cavalli-Sforza, 1970; Eysenck & Kamin, 1981). These arguments beg the question as to what constitutes intelligence in the first place (Chase, 1980; Sternberg, 1988).

The second category of theory explained the white-minority achievement gap as a function of cultural deprivation. Minorities simply do not have the kind of cultural capital that is necessary for success in a white-dominated society. This was suggested

to be particularly true of urban blacks. Referencing Lemann (1986), Sizemore asserted that a major problem of inner city blacks is that they have been abandoned by the best of their cultural peers. Members of the black working and middle classes, freed from housing discrimination by the civil rights movement, fled the ghettos taking with them their "ethic of assimilation." Bereft of their leaders and their institutions, remaining blacks fell into "a kind of free fall" (p. 43).

Sizemore suggested that a third category of theory imputes to the black family and community qualities of indifference and immorality. Divorce and separation, teenage pregnancy, and nonemployment and unemployment are said to contribute to families which are unable to provide for the educational (or emotional) needs of their young (McGhee, 1985). Sizemore asserted that these "are the most serious of all problems faced by Black families because they are the primary reason for their poverty (p. 43)." Dysfunctional structural conditions are suggested to be a dominant force in the decline of the black nuclear family: "The basis of a stable family rests on the willingness and ability of men and women to marry, bear and rear children and fulfill socially prescribed familial roles (Staples, 1987, pp. 277-78)." Sizemore claimed that data showed that, while black women were willing to fulfill their roles as wife and mother, black men were thwarted in their efforts to be economic providers and family leaders. Unemployment and nonemployment were said to be the chief causes (pp. 44-45).

Sizemore was not nearly as expansive in her exposition of the remaining two theories which explain achievement disadvantage as a function of (4) underfunded, ineffective schools and (5) schools' perpetuation of a racial/caste system. As a counter to ineffective schools, Sizemore alluded to the body of Effective Schools literature suggesting that all children can learn given sufficient attention to the characteristics associated with schools in which disaggregated achievement data demonstrated ethnic,

socio-economic status, and gender equity. Perhaps Sizemore's concluding comments best illustrate her racial/caste system theory:

We have the technology to rescue at-risk learners from failure. It is no longer an educational question. It is a political problem. Politics is the management of the conflict of groups at war over scarce resources. Power is the central concept. When those who have power decide that we should educate the at-risk students, we will do so on a grand scale (p. 51).

Sizemore thus extended earlier environmental theory to include the entire society. The remainder of her writing in this chapter addressed the black experience in the United States and outlined steps to sensitize mainstream America, and particularly black school children, to its historical importance.

Although Sizemore (1990) may be criticized for her cursory treatment of the last two categories of theory, like Lansa and Potter, she may be credited with attempting to provide a multi-dimensional, historical view of why minority students seem to have difficulty achieving at levels comparable to their white peers. Like both Lansa and Potter (1984) and Haycock and Navarro (1988), Sizemore theorized that the environment is a causative factor in the white-minority achievement gap. (One conjectures that the Jenson (1969) theory is cited only for its historical significance.) Moreover, in citing the importance of the schooling experience and the primacy of expectations, Sizemore aligned herself with the earlier theorists. Unlike the previous frameworks, however, the capstone of Sizemore's model (the racial/caste system) indicted the entire social fabric, suggesting that the inequities of school achievement are a reflection of the inequities of the larger society. Writers such as Kozol (1992) echoed this belief. An outline of the theories crossing authors with explanations is illustrated in Table 1.

Table 1

Outline of Theories Explaining the White-Minority Achievement Gap

	<u>Lansa and Potter</u>	<u>Haycock and Navarro</u>	<u>Sizemore</u>
<u>Explanation</u>			
1. Race membership	Jenson's genetic		Jenson's genetic
2. Home, family, and community	Coleman's background	Home and family	Cultural deprivation
	Cultural deprivation		Dysfunctional family and community
3. Home and school interaction	Cultural-difference		
4. School	Effective Schools	Schooling experience	Effective Schools
5. Society and power structures			Racial/caste System

Beyond their theoretical importance, however, explanations for the existence of the white-minority achievement gap may serve also as organizers in examining attempts by educators to reduce the gap. The next section reviews some of these attempts.

#### What Are Some Attempts to Reduce the White-Minority Achievement Gap?

Seen as the application of theory, a large body of literature examines programmatic attempts to reduce the white-minority achievement gap. Not surprisingly, much of this literature is written by education practitioners and researchers in the states experiencing the most dramatic increases in language-minority student enrollment. Alternately, other literature reports on programs targeting primarily inner-city black students within major U. S. urban centers.

Acknowledging that problem solution is based upon problem definition (Achilles, 1987), this writer will attempt to review the related literature of this section using the theories examined in the previous one. In two cases, the citation of the theory-problem is the organizer; in a third, it is the theory-solution that serves as the focal point. There is a sense in which the theories may be viewed as a continuum: one end represents those who would attribute school success as a function of the home and family, the middle represents those who believe that the interaction of the home and school is the agentic factor, and the other end represents those who hold the school and its practices as uniquely responsible for student achievement. The following section illustrates those writing from the first perspective.

#### Cultural Deprivation as the Problem

Cultural deprivation theory (Lansa & Potter, 1984; Sizemore, 1990) is based on the premise that deficient home and community conditions of minority youngsters precipitate school failure. When minority children are removed from their communities and schooled with more advantaged white children, an environment conducive to higher achievement levels for minority children is created. School desegregation and forced busing are examples of the application of cultural deprivation theory. Since the Jencks et al. (1972) re-analysis of the Coleman et al. (1966) data, numerous reports have emerged evaluating the effect of school desegregation and/or forced busing on the reduction of the white-minority achievement gap.

Writing for the National Association for Neighborhood Schools, an avowedly anti-busing organization, D'Onofrio (1983) analyzed student standardized norm-referenced test (NRT) scores before and after busing in New Castle County (DE) to determine the effect of forced busing on white-black achievement gap reduction. In contrast to local education agency (LEA) evaluators who utilized as baseline data students' scores on the California Achievement Test (CAT) from the first year of busing

(1978-1979) , D'Onofrio obtained data from test administrations four years prior and hence to the first year of busing and used those as "pre- test" scores. For example, D'Onofrio reported that in grade 1 for 1981-1982 (the fourth year of busing), the gap between whites and blacks was 26.2 points, double that of the "segregated" years (p. 7). His findings suggested that the differences between white and black scores on standardized tests of achievement were greater after forced busing for desegregation than before. In terms of closing the achievement gap, he concluded that desegregation was not working: "forced busing in New Castle County has been an abject failure (p. 8)."

The D'Onofrio study, however, may be criticized on a number of counts: Standardized achievement tests utilized by the LEA prior to the first year of busing included two tests (the Cooperative Primary Test and the Sequential Test of Educational Progress) which were statistically correlated neither with each other nor the CAT used in post-busing years. Accordingly, the extent to which D'Onofrio is "mixing apples and oranges" is left to conjecture. Moreover, the test scores in the pre-busing years were not disaggregated by race. D'Onofrio assumes that, by comparing the scores of predominantly black with predominantly white schools, he is obtaining measures of achievement which are uniquely attributable to race. Reasonably, other factors (e.g., as economic advantagement) could as easily account for observed differences as racial membership. One must also question the objectivity of a report "devoid of esoteric techniques and jargon" written by an "insurance agent and tax accountant" serving the interests of "the foremost national anti-busing citizens' lobby" which repeatedly refers to the LEA's desegregation plan as a "scheme" (p. 6). At best, the D'Onofrio report is viewed as biased; at worst, the findings are altogether invalid.

Three studies addressed desegregation and its effects on narrowing the white-minority achievement gap in the Cleveland (OH) Public Schools ( Zafirau, 1986;



Lanese, 1988; Gallagher & Lanese, 1991). Using as the criterion measure reading comprehension scores from the California Test of Basic Skills (CTBS), Zafirau (1986) analyzed the progress from 1982 to 1986 of students adversely affected by prior racial segregation in the Cleveland (OH) Public Schools (blacks) and compared their progress with the non-adversely affected students in the district (whites). "Only if their [blacks'] progress is greater than for the non-adversely affected students will the discrepancy between the two groups be reduced (p. 2)." Zafirau used regression analysis to predict non-adversely affected students' posttest scores from knowledge of their pretest scores. To model the results that might have been obtained had blacks been part of the non-adversely affected group, the same regression equation was then used to predict posttest scores of adversely affected students. Comparing the predicted scores to the actual posttest scores for each student in the adversely affected group allowed Zafirau to determine whether the adversely affected group's actual progress in reading comprehension achievement was the same, worse, or better than it would have been had it progressed at the same rate as the non-adversely affected group. Citing elementary school data, Zifirau reported a 3.62 NCE improvement from predicted to actual score; similar gains were reported for the junior and high school level (p. 14). Zafirau concluded: "Although the reading comprehension achievement gap between the adversely affected and non-adversely affected student has narrowed from spring of 1982 to spring of 1985 (system-wide and for every type of school), the remaining achievement gap between these two student groups is still educationally significant (p. 18)." Accordingly, Zafirau suggests that desegregation in the Cleveland (OH) Public Schools has had limited but beneficial results in closing the white-minority achievement gap.

Zafirau's study leaves little to criticize. Unlike the D'Onofrio (1983) study, the Zafirau report (a) demonstrates the use of statistically valid comparisons and (b)

controls for race as a variable. Consequently, threats to validity are minimized. Moreover, (c) there is no evidence in the language of the study to suggest that the investigator has "an ax to grind."

Two other reports (Lanese, 1988; Gallagher & Lanese, 1991) of progress toward attaining parity in the CAT reading comprehension scores of white students and black students in the Cleveland (OH) Public Schools have been published. Both longitudinal studies proceeded from a court order which mandated desegregation of the Cleveland schools in 1980-1981 and the implementation of the Affirmative Reading Skills Program in 1982-1983. Parity was defined as "statistically equivalent proportions of black and white students scoring at or above the thirty-fourth percentile rank on the reading comprehension section of a NRT." Data for the annual multi-year cross-section analysis and a longitudinal analysis were compiled. Examining the test scores that were available for 12 cohorts, ranging in size from 1400 to 2900 students, Lanese (1988) discussed the findings for the years 1979-1987. The following results were reported: (1) the difference between the percentage of white students and the percentage of black students scoring above the thirty-third percentile widened over the course of the study for all cohorts spanning more than two years; (2) parity of test scores did not occur in any of the cohorts in the final two years of the study; (3) the parity gap is lower in the elementary grades for all cohorts and parity occurs more frequently in the lower grades; (4) black students read less well than white students at most grade levels in each of the 12 cohorts.

Gallagher and Lanese (1991) reported on progress toward parity for the following two years, 1988-1989 and 1989-1990. Results pertinent to the present study indicated that: (1) racial parity in reading results was attained for the first and second grades in 1989-90; (2) parity gaps for secondary school grades have decreased since the onset of desegregation; (3) longitudinal analysis, following cohorts of

students, indicated increasing parity gap by grade; and (4) individual schools had widely varying success in improving reading comprehension at various grades.

Both Lanese (1988) and Gallagher and Lanese (1991) epitomize at the district-level that which is demonstrated by the nation's schools as a whole--there is a white-minority achievement gap and it grows proportionally wider in relation to grade level. However, desegregation, or perhaps more accurately, the district's compensatory programs implemented since desegregation ("Major Works," THINK, and STAR), appeared to be effective in reducing the white-minority achievement gap in reading comprehension; however, the benefits did not accrue uniformly across all schools, all grades, and all student cohorts. This finding is seen to corroborate Zifirau's (1986) conclusion that desegregation in the Cleveland (OH) Public Schools had limited but beneficial results in closing the white-minority achievement gap.

Mitchell (1989) evaluated a state-financed school desegregation plan in metropolitan Milwaukee (WI). The basic assumption of the plan was that with enough time, state financial assistance, and new spending, racial integration would significantly improve academic achievement and close the gap between minority and White academic achievement. Mitchell reported that since its court-mandated implementation in 1976, racial desegregation through forced busing has occurred at 112 of the Milwaukee Public School (MPS) schools. Moreover, the pupil assignment system has contributed to a pattern of student segregation along economic class lines such that students from families with less income, less stability, and parental involvement tend to attend traditional MPS schools while students from families with greater income and stability more likely attend specialty MPS schools and suburban schools. Based on descriptive analyses, Mitchell found the following with regard to academic achievement: (1) Seventy-eight percent of black MPS students attend traditional, non-specialty schools; (2) Average grades for black students were D to D+;

(3) Between 26% and 43% of black grades were F; (4) The percent of blacks exceeding national test averages ranged from 8% to 21%; (5) Significant gaps existed in black and white achievement; in some cases the gap is closed in part because of declining White achievement; (6) Twenty-two percent of black MPS students attend city specialty schools or suburban schools. Their average level of academic achievement exceeds that of blacks in traditional MPS schools. Significant gaps in black-white achievement also exist in specialty and suburban schools (p. 3). Mitchell concluded that desegregation has "failed as a strategy for increasing academic achievement of black students" and instead "substituted a system of segregation by economic class for one based on race." Moreover, Mitchell asserted that busing for integration has "prevented other strategies from being used which would have greater probability of increasing academic achievement and has placed an inequitable and discriminatory transportation burden on black families, reflecting a conscious policy to minimize white opposition to busing (p. 4)."

Writing for the Wisconsin Policy Research Institute whose stated goal is "to provide nonpartisan research on key issues that affect citizens living in Wisconsin (p. 114)," Mitchell is persuasive. Although later sections of his report tend toward polemics, Mitchell must be credited with compiling a prodigious quantity of data to support his conclusions. As with the less creditable D'Onofrio (1983) research, Mitchell's study argues against the efficacy of desegregation (or at least forced busing) as a means to close the white-minority achievement gap. (Interestingly, in a recent article published by Education Week, Wisconsin State Superintendent of Public Instruction Herbert J. Grover referred to the Wisconsin Policy Institute as "a political hatchet operation for the right wingers (DiegmueLLer, 1993, p. 1)" thus casting some doubt on the extent of the Institute's commitment to 'nonpartisan research.')

To evaluate the 1986 court-mandated busing policy in the Norfolk (VA) Public Schools, Ipka (1992) examined the extent to which individual characteristics, school characteristics, and busing affected the white-minority achievement gap between the busing and postbusing years. His methodology involved multiple regression analysis of the achievement test scores of 228 African-American and 203 white fourth-grade students for the busing year 1985-1986 and the nonbusing year 1986-1987 (the dependent variable) and individual and school characteristics (the dependent variables). Ipka's findings indicated that positive relationships existed between the gap in achievement test scores and Chapter 1, race, school income, and the average number of library books. Negative relationships were found to exist between the achievement gap and gender, school building age, and average teacher salary. Ipka stated: "Race as a variable affecting the achievement gap between the busing and nonbusing year also appeared to be statistically significant ( $p \leq .05$ ) in explaining the variance in test scores of students. The T value 2.24 indicated that as the number of white students increased in a given school, the achievement gap decreased (p. 29)." Although Ipka argued that conclusive statements could not be generated from his analysis, to the extent that lower achieving minority students appear to achieve more when schooled with higher achieving white students (or more specifically, enjoy the economic advantages of being in a "white" school), her research supports desegregation as a solution to reducing the white-minority achievement gap.

As in the Zifirau (1986), Lanese (1988), and Gallagher and Lanese (1991) studies, neither methodological nor conceptual criticisms seem warranted regarding the Ipka (1992) report. Although exemplifying associational field (as opposed to experimental) research, her sample is both representative and numerically adequate and her statistical procedures sound. Moreover, Ipka does not hypothesize beyond her data, instead remaining conservative and cautionary in her conclusions.

A preponderance of the evidence, based on the preceding studies of primarily urban communities, suggests that desegregation as a solution to reducing the white-minority achievement gap may be effective. However, one must note the caveats associated with desegregation: (1) Public confidence in the schools may abate when initial gains in minority achievement are not attained (D'Onofrio, 1983); (2) Schools formerly segregated by race may become resegregated by economic class as advantaged people, whites and blacks alike, abandon traditional urban schools for suburban or magnet schools (Mitchell, 1989). Essentially, these observations are but reiterations of the Armor (1972), St. John (1975), and Coleman (1975) research cited in a previous section of this chapter. Predating all but one of the local education agency studies, however, was a study utilizing national databases. This study (Jones, 1984) provides a foundation to explain both the positive and negative findings seen in the studies reviewed above.

Writing for American Psychologist, Jones (1984) marshaled evidence to suggest that the narrowing of the gap between the scholastic achievement levels of white and black youth demonstrated since 1971 may be attributable, in part, to school desegregation. Jones examined the relative changes in students' (a) reading and mathematics performance on the National Assessment of Educational Progress (NAEP) and (b) verbal (SAT-V) and mathematics (SAT-M) scores on the Scholastic Aptitude Test (SAT). The NAEP data were the result of test administrations to 9-, 13-, and 17-year olds in 1971, 1975, and 1980. The reading section of the test assessed students' abilities to comprehend, analyze, use, reason from, and make judgments about what they have read. At each assessment year, the same group of reading exercises was administered to a national sample ( $n \geq 2,000$ ) of students at each target age. From these data, Jones plotted differences in average performance (percentage of items correct) of white students and black students for the three assessments by year of

birth, 1955, 1960, 1965, and 1970, thus producing trend lines. This procedure provided trend lines not only at age 13 but also at ages 17 and 9. The trend lines showed a dramatic decline in the white-black average difference at each age: The white-black average difference of 20% correct for the birth year of 1953 was reduced to 10% correct for the birth year of 1970. Utilizing the same procedure to examine average differences in SAT-V scores, Jones obtained a similar result: There was a steady decline in the average white-black difference: During the period 1976 to 1983, average SAT-V scores for white students declined by 8 score points, but average scores for black students increased by 7 score points.

Jones then examined relative changes in quantitative skills. The NAEP exercises were developed to assess knowledge, skills, understanding, and application in several content areas: numbers and numeration, variables and relationships, geometry, probability and statistics, and graphs and tables. By plotting the white-black average differences in percentage correct against children's year of birth, Jones illustrated a consistent decline in the average difference between whites and blacks at each age and significantly higher relative levels of average performance in mathematics for black children born since 1965 than for black children born earlier. Similarly, when he examined SAT-M data plotted by year of birth, Jones found that white-black average differences consistently declined: During the period 1976 to 1983, average SAT-M scores for white students declined by 9 score points, but average scores for black students increased by 15 score points. Jones stated that similar analyses of NAEP data for other subjects--writing, science, and social studies--supported the same conclusions for achievement in those areas (Burton & Jones, 1982; Hueftle, Rakow, & Welch, 1983).

Jones concluded:

The evidence for a narrowing gap between average scholastic achievement levels of white and black youth seems incontrovertible. However, we cannot with confidence attribute the narrowing gap to educational or social programs that were initiated with this as their goal. It is possible that the trend reflects long-term effects of school desegregation, even though such effects have proven elusive to investigation in shorter time spans (e.g., Gerard, 1983; Gerard & Miller, 1975). Perhaps school desegregation failed during its early years, often marked by tension and strife, but more recently has had beneficial educational effects for minority students. Or perhaps, as concluded by Crain, Mahard, and Narot (1982), achievement gains are seen for black students attending desegregated schools from kindergarten or first grade, but not for black students whose desegregated education began in later school years. Evidence consistent with either of these possibilities is provided by Moe, Nacoste, and Insko (1981). They assessed student attitudes about students of different ethnic origin (black or white) in the same small-town North Carolina junior high school in 1966, a year after the school was integrated, and again in 1979, when the students had spent their entire school experience in integrated classrooms. White students in 1979 were far more accepting of black students in a wide variety of personal and social interactions than white students had been in 1966. To the extent that this finding would apply in other schools, such changes in attitude may create a school environment more conducive to higher achievement levels for black students (p. 1212). (Emphasis added.)

Based on his analyses of the NAEP and SAT data and from his reading of the literature, Jones thus suggested three theories relative to optimizing the effects of school desegregation on the white-black achievement gap: (1) Desegregation narrows the gap after the social upheaval that accompanies its implementation abates; (2) Desegregation narrows the gap when its implementation includes the primary years of school; and (3) Desegregation narrows the gap when whites' acceptance of blacks creates an environment conducive to black student achievement.

Jones' (1984) contribution to Cultural Deprivation as an explanation of the white-minority achievement gap, and specifically to desegregation as agentic in reducing the gap, is immense. Given that "1969-1970 marked the largest wave of court-ordered school desegregation in the South (p. 1207)" and that the majority of blacks reside in the South (Hodgkinson, 1992), Jones' utilization of year of birth as a



basis for plotting NRT data provides a means by which the interaction of the implementation of desegregation with student age can be assessed. Jones' theoretical contribution must also be recognized as deepening our understanding of the dynamics of desegregation as an intervention. For example, to the extent that D'Onofrio's (1983) findings arguing against the efficacy of desegregation are accurate, they may be explained by initially hostile attitudes mitigating against the acceptance of blacks in formerly white schools. Jones theorizes that without social acceptance from the white majority, black achievement is difficult. Seen in the larger context of Cultural Deprivation theory and most pertinent to the present research, one of Jones' theories suggests that if the achievement gap is to be reduced, the cultural deficiencies of minority homes must be counterbalanced by early and prolonged integration with advantaged majority peers.

In conclusion, the literature reviewed in this section argues persuasively for desegregation as a solution when Cultural Deprivation is viewed as the problem. Although forced busing to achieve desegregation may generate ancillary problems, one may conclude that, given sufficient time and acceptance from the white majority, achievement gains accrue to disadvantaged minority students when they are schooled with more advantaged white children. Moreover, the minority student gains do not seem to be at the expense of white student achievement (Jones, 1984, p. 1208). Finally, the literature suggests that desegregation optimally narrows the achievement gap when it includes the primary years of school. The next section reviews literature written from the perspective of Cultural-Difference theorists.

#### Cultural-Difference as the Problem

Lansa and Potter (1984) stated that cultural-difference theorists rejected the views of the cultural deprivationists; rather they argued that the culture of the school

must be reformed commensurate with the diverse cultures of minority students. Cultural-difference proponents held that neither home nor school exclusively accounts for student achievement, that instead the interaction of home and school is the controlling factor (p. 7). Multicultural education, English as a Second Language (ESL) and Limited English-Proficient (LEP) programs are the programmatic application of cultural-difference theory.

In a report on California's experience with LEP Hispanic students, Gold (1985) argued that the competency testing movement served to widen rather than to narrow the achievement gap between minority and non minority students. Admitting that the movement gave cohesion to the curriculum, guided scarce resources for remediation, motivated students, teachers, and parents to higher performance standards, and increased accountability, Gold nevertheless suggested that major problems were also created. He argued that greater alienation of LEP students from school, more dropouts, and inadequate development in academic and vocational skills resulted from increased course requirements, greater reliance on tests in English, and higher expectations of skills demonstrated only through English. Gold recommended that educators: (1) Develop competency tests at grades 3, 6, and 8 in several languages; (2) Avoid grade retention whenever possible; (3) Communicate well to LEP students and their parents diploma sanctions and advise them of their options; (4) Use the primary language in newsletters and bulletins to provide competency test information to parents and community members; (5) Ensure psychometric rigor in the local- and state-mandated tests to define clearly competencies, correct devised items, and to ascertain reliability/validity; (6) Discern between the need for specific skill remediation versus continued language development; and (7) Differentiate between the language of instruction and of remediation.

Although Gold (1985) presumably draws upon empirical bases in formulating his observations, his report is best categorized as a policy brief, written for the purpose of influencing education decision-makers (and ultimately practitioners) as opposed to contributing to the research base. Specific research studies to support or refute Gold's recommendations are not cited. Nevertheless, in illuminating the link between competency testing and its possible effect on the achievement gap between Anglos and LEP minority students, Gold's report is informative to the present study.

Drawing upon demographic studies and the Effective Teaching literature, Schuhman (1987) suggested various ways of improving teacher quality to help close the achievement gap between majority and minority students. Prior to citing specific recommendations, Schuhman cited an alarming trend: In 1970 black professionals made up 12% of all teachers at the elementary and secondary levels, while today they constitute about 8% of that pool (NCES, 1983). Hispanics are reported to be even more seriously underrepresented. Schuhman wrote: "The absence of a representative number of minority teachers and administrators in a pluralistic society distorts social reality for children and is detrimental to all students, white as well as minority (p. 153)." Schuhman suggested that effective teachers of minority students (a) exhibit active teaching behaviors, (b) communicate high expectations for learning, (c) have a knowledge and appreciation of their pupils' culture and use this knowledge for instructional purposes, and (d) recognize the legitimacy of the language variety of the students and use the students' language in developing English (p. 158).

Schuhman (1987) claimed that her report is based both on the Effective Teaching research and on her experience with minority students. As with the Gold (1985) paper, Shuhman's report appears oriented toward influencing education decision-makers rather than augmenting the research base. However, to the extent that the closing of the white-minority achievement gap is dependent upon teacher behaviors

and the policies that affect them, Schuhman makes a contribution to the present research effort.

In a report from the Honolulu (HI) School District, Yap (1988) evaluated the district's Students of Limited English Proficiency (SLEP) program and its relative effect on reducing the achievement gap between the English-speaking majority and LEP minority students. Specifically, the evaluation addressed whether the program was meeting its objectives, whether ESL instruction, as contrasted with bilingual instruction, should be continued, whether the program's pull-out design should be continued, and how program costs relate to outcomes. Using a pretest/posttest design, SLEP scores on the Metropolitan Achievement Test (MAT) and the Basic Inventory of Natural Language (BINL), were compared and analyzed. The findings showed that the SLEP program participants were making achievement gains in oral English, reading, language arts and mathematics beyond what is expected of comparable students in the regular classroom, as measured by standardized NRTs (MAT and BINL). Not only was the achievement gap between LEP students and non-LEP students narrowed but, in many cases, the growth rate of comparable nationally normed groups was surpassed. Yap also reported that, as a group, the SLEP participants were doing satisfactory or better work in their regular classroom, as measured by grade point average (GPA) with a failure rate of five percent or less in reading, language arts, mathematics and speaking/listening. Yap concluded that her research supported the popular use of the pull-out setting and ESL instruction in the SLEP program from a cost and effect standpoint (p. 3-4).

Yap's (1988) work is methodologically sound. Her instrumentation, research design, sample selection, and statistical procedures provide for findings that are verifiable and subject to replication. Yap's contribution to the present study of white-minority achievement gap reduction lies in the demonstration that intervention

specifically targeting LEP students as a minority population can make a difference. Moreover, the success of the pull-out design of the program evaluated by Yap suggests that reduced class-size is an effective strategy to close the achievement gap.

Parrenas and Parrenas (1990) reviewed the meta-analyses and literature on cooperative learning (CL) to determine the effectiveness of CL as a tool for closing the achievement gap between traditional majority students and the "new majority" (LEP minority group students whose numbers equal or exceed the old majority). Slavin (1983) was cited as a major source. Of 46 controlled studies of elementary and secondary classrooms, Slavin reported that 63% showed superior outcomes for CL, 33% showed no differences, and only 4% showed higher achievement for the traditional comparison groups. Almost all (89%) of the studies which used group rewards for individual achievement (individual accountability) showed that achievement was about the same as in comparison classrooms. "The lowest achieving students and minority students in general benefited most, but the benefit obtained for the lower achievers is not bought at the expense of the higher achievers; the high achieving students generally perform as well or better in cooperative classrooms than they do in traditional classrooms (Parrenas & Parrenas, p. 5-6)." Parrenas and Parrenas stated:

Considerable research demonstrates that minority students are relatively more cooperative than majority students in their social orientation. Thus, the particularly strong gains of minority students in cooperative learning classrooms may be due to the compatibility of the classroom structure with the individual social values of minority students. The choice of exclusively competitive and individualistic classroom structures may bias academic and social outcomes against the achievement of minority students (p.17).

Citing positive interdependence, individual accountability, group processing, interpersonal and small group skills, and face-to-face promotive interaction as basic elements of CL, Parrenas and Parrenas suggested that CL can help eliminate ethnocentrism and racism even as it enhances achievement gains.

As a review of the literature, the work of Parrenas and Parrenas (1990) is far from exhaustive. (Only nine sources are listed as references.) The authors seem to rely on the meta-analyses of others (e.g., Slavin, 1983) as opposed to undertaking a meta-analysis of their own. In and of itself, this is not a criticism. However, there is a body of research not mentioned by the authors that suggests that the effectiveness of CL is conditional and context-specific (Slavin, 1990). Nevertheless, Parrenas and Parrenas' work suggests another possible strategy that might be employed in narrowing the white-minority achievement gap. Moreover, to the extent that CL can be considered as a technique for reducing class-size (most CL groups are four-student teams), the reported effectiveness of CL in enhancing minority achievement provides support for the major thesis of the current research.

Writing for the National Center for Research on Cultural Diversity and Second Language Learning, Garcia (1991) reviewed the research on effective instructional practices, revealing numerous programs and practices that tended to reduce the achievement gap between the English-speaking majority and LEP minority students. Garcia's report was based primarily on case studies documenting educationally effective practices in California and Arizona: Carpenteria (Cummins, 1986); San Diego (Carter & Chatfield, 1986); the San Francisco Bay area (Lucas, Henzel, & Donato, 1990; Pease-Alvarez, Garcia & Espinosa, in press); Phoenix (Garcia, 1988; Moll, 1988). Garcia found that the effective instructional practices employed (a) high levels of communication, (b) small ( $n \leq 8$ ) groups of students, (c) integrated and thematic curriculum, (d) collaborative learning, (e) language and literacy, and (f) awareness of cultural perceptions. Regarding (b) and (d), Garcia reported that "Teachers consistently organized instruction so as to insure heterogeneous small-group collaborative academic activities requiring a high degree of student-to-student interaction (p. 7)." Garcia concluded that linguistically and culturally diverse students

can be served effectively, that they can achieve academically at levels at or above national norms, and that instructional strategies that work best acknowledge, respect, and build upon the language and culture of the home (p. 16).

Like the Parrenas and Parrenas (1990) report, Garcia's (1991) work is a synthesis of research focusing on the achievement of minority students. Rather than examining the effects of a particular strategy (e.g., cooperative learning), Garcia, like Schuhman (1987), looked at schools where instructional practices resulted in narrowing the achievement gap between LEP students and the Anglo majority. In citing the importance of cooperative (collaborative) learning, small groups, and appreciation for cultural diversity, Garcia's review corroborates previous work.

In conclusion, the literature reviewed in this section demonstrates a high degree of consensus among those writing from the cultural-difference perspective. The literature suggests that the interventions most effective in reducing the achievement gap are those that acknowledge and capitalize upon the differences between home and school. As an explanation of the gap where primary cultural and linguistic differences exist (Ogbu, 1992), the cultural-difference perspective seems particularly compelling. In a continuum of causality, cultural-difference theory may be seen as a midway point, bridging cultural deprivation theory, the belief that conditions of home and family are primarily responsible for student achievement, with Effective Schools literature, which argues that student academic success or failure lies with the school. The next section focuses on the writing of educators whose work (primarily field research reports) demonstrates the attitudes and practices of Effective Schools as a solution to the white-minority achievement gap.

#### Effective Schools as the Solution

Building from the work of Ron Edmonds (1979a) and Larry Lezotte (1989), Effective Schools is as much a process as a theory. Nonetheless, in their reviews, Lansa

and Potter (1984), Haycock and Navarro (1988), and Sizemore (1990) uniformly accommodated the Effective Schools' beliefs that schools do make a difference and all students can learn. Similarly, each reviewed the school programs and practices (correlates) that have been found to be associated with schools where the achievement gap between white and minority students is closing. (Additional Effective Schools literature of both theoretical and practical significance includes Brookover et al., 1979; Edmonds, 1979b, 1981, 1982; Levine and Lezotte, 1990; Peterson and McCue, 1990.)

In an Interim Evaluation Report for the Danforth Foundation and elsewhere, Achilles and Duvall (1982, 1983) reported on the implementation of Project SHAL, an Effective Schools program in four St. Louis (MO) public schools (Stowe, Hempstead, Arlington, and Laclede). The authors stated that "Project SHAL is essentially a pilot effort to determine if key efforts based on the 'effective schools' research would have a substantial impact in raising test scores of youngsters in four schools (1982, p. 1)." Described as typical inner-city schools where the pupils were performing well below norm on standardized tests, the schools were "nearly 100% minority (1983, p. 1)." Accordingly, a goal of Project SHAL was to assist students to attain and maintain the national norm (CAT scores) on standardized tests in the basic skills. Project SHAL utilized inservice programming that focused on teachers, parents, and others affiliated with the schools. Program implementers attended to five major elements of effective schools (building leadership, high expectations, focus on basic skills, a school climate conducive to learning, and frequent monitoring of pupil progress). Inservice efforts followed the (a) Teacher Expectations and Student Achievement (TESA) and (b) Expectations, Interactions, Achievement (EIA) program models. Using pre/post comparisons from the original four schools (12 additional schools were added in the second year of the project), SHAL schools showed "gains" (defined as increased in the



percents of pupils in quartiles 3 and 4 and reductions in the percents of pupils in quartiles 1 and 2) in math and reading. Statistically, 62% of the total test results of the SHAL schools attained or exceeded the national mean, and in all but one case, each SHAL school was closer to the National Norm than the City-wide result.

Although the Achilles and DuVall (1982, 1983) reports were only preliminary, the initial findings suggest that NRT scores of inner-city minorities can be positively affected by attending to Effective Schools correlates, where such attention includes district leadership in specific staff development activities and systemic monitoring of NRT results. In effect, by increasing the number of minority students in the upper quartiles and reducing those in the lower ones, the white-minority achievement gap is narrowed.

In a paper presented at the annual meeting of the American Association of School Administrators, Sang (1987) reported on the Duval County (FL) Public Schools' efforts to reduce the gap between black and white students' academic achievement. Interventions were based on the "Mayo Clinic Approach" to education involving "a team approach to not only treat the symptoms, but address the causes of those symptoms. Just as all patients are treated individually regardless of their similar symptoms, students are treated as individuals with unique circumstances (p. 1)." Based on the results of an unpublished "Gap Study" (n.d.) showing wide disparities in the NRT scores, study skills, attitude, thinking and listening skills, and test-taking strategies of black versus white students, Duval County educators developed and implemented several programs: (1) The Stimulating Aptitude Skills (SAS) project, targeting black high school students interested in taking the Scholastic Aptitude Test (SAT), purported to increase students' test taking skills. (2) The Teachers As Interpreters program sought to provide teachers with the ability to interpret a student's scores, pinpoint deficiencies, and implement more effective teaching strategies. (3) Project BASE

(Blacks for Academic Success in Education), involving local businesses and churches, identified and publicly celebrated black students in grades five through eleven with scores above the 85th percentile on the reading and/or mathematics subtests of the SAT. (4) Lifestyle changes were encouraged through (a) open houses, (b) summer job placement, (c) homework assistance (Homework Hotline and The Learning Hour), and (d) grief therapy. (5) The "Make a Difference" inservice program assisted teachers in raising their expectations for minority students. (6) Classroom observation with a peer partner taught teachers to code teaching behaviors, try new teaching behaviors, and implement more successful teaching strategies.

*Reported results included:* (1) Black students' average SAT reading scores rose from the 23.4 percentile in 1976 to the 45.6 percentile in 1986, a gain of 94.9%. (2) Black students' average SAT mathematics scores rose from the 25.3 percentile to the 51.8 percentile, a 104.7% gain, in the period 1976-1986. (3) The number of black juniors taking the Preliminary Scholastic Achievement Test (PSAT) increased from 176 in 1981-1982 to 410 in 1985-1986; the number of black sophomores taking the PSAT increased by 1,200% during the same period. (4) Enrollment in Advanced Placement (AP) courses by black students increased by 129%. (No time period reported.) (5) Grade point averages (GPA) increased. (No time period reported.) (6) Fifty percent of the BASE students surveyed showed improved attitudes toward school work. (7) The number of discipline problems among BASE students decreased. (No time period reported.) (8) The number of students qualified to join Project BASE increased each year between 16% and 30%. (9) The gap between white students' NRT scores and black students' NRT scores decreased by seven points.

The Sang (1987) report displays a number of deficiencies including the lack of sufficient baseline data for pre- and post-intervention comparison for AP courses taken, GPA, and discipline reports. Similarly, the author fails to describe procedures

for obtaining the attitude survey sample. Moreover, no results of the teacher-oriented programs are reported. Despite researchers' inability to fully control a multitude of variables in field research (of which Sang is an example), one would hope for more complete disclosure in a report of how and when data were obtained and the results drawn from them. Nevertheless, Sang (1987) is helpful in demonstrating the possible benefits of launching an assault on the white-minority achievement gap from many fronts at the same time. Commensurate with Effective Schools correlates, high expectations from both the teachers and the larger community are seen as an integral component of successful intervention.

Alarmed by the publication of a 25-point CAT score gap between white and minority students in a Maryland elementary school, school superintendent John Murphy (1988) launched (and subsequently reported on) an Effective Schools' process in 171 schools. Murphy said the first step was to confront the fact that a "gap" existed and to allow it to be a catalyst for change. Convened by the superintendent, a task force of community leaders and educational experts found that "the root cause of the gap was negative attitudes about the potential of black students (p. 41)." A subsequent action plan emphasized the academic achievement of black students, provided for higher expectations, added resources, and a stronger core curriculum. Specific interventions included: (1) altering existing efforts in reading, mathematics, and writing to foster minority participation and performance; (2) a program to increase black and Hispanic students' participation and achievement in mathematics and science; (3) staff training in the Teacher Expectations and Student Achievement program (TESA) and the Improving Minority Student Achievement program from American University's Equity Center; (4) implementing an Effective Schools Process district-wide; (5) strengthening curricula and emphasizing fundamentals, including student discipline, dress codes, and punctuality; (6) implementing the Millikan II program which fostered

high expectations for all children through (a) parental involvement, (b) team planning, (c) a positive social and academic environment, (d) adapting the school system's curriculum to specific needs of students, (e) reducing staff-to-student ratios, and (f) providing extra instructional materials and computers; and (7) implementing Project SUCCESS which focused on at-risk students in the high schools. Murphy reported gains on nearly every measure and across every intervention program. For example, from 1985 to 1987, Millikan II schools' gains on the CAT outdistanced district gains by almost two to one; black students who attended the highly structured Traditional/ Classic Academy magnet program realized "substantial gains" on the CAT, a 30% increase in one year (p. 41).

Murphy (1988) is more "press release" than research report. As such the validity of specific findings is questionable. One can only hypothesize, for example, as to the extent of the Hawthorne effect given such major attention brought by the charismatic Superintendent Murphy to the "gap problem" in Prince Georges County. As Sang (1987), Murphy nonetheless demonstrates the all-out assault that school district leaders can wage on reducing the White-minority achievement gap once the will to do so has been publicly established. Commensurate with all Effective Schools efforts, the primary beliefs were explicit--all students can learn and schools are responsible for students' learning.

Denton and Davis (1988) described programs believed to be successful in reducing the White-minority achievement gap in the Pittsburg (PA) Public Schools. The authors stated that the initial enthusiasm generated by district school improvement in the early 1980s was tempered by the results of the disaggregation by race of the CAT student scores. The test scores indicated that "the increases in achievement by white students in each content area far surpassed the achievement gains by black students. . . The disaggregated analysis of achievement scores further revealed that the achievement

gap tended to widen across grade levels (p. 2)." Consequently, the Pittsburg Public Schools Board of Education identified and implemented specific programs to reduce the achievement gap. Denton and Davis reported that several interventions were district-based: (1) Community Educational Activity Centers, (2) Explorers of Growth Opportunities, (3) Pittsburg Achievement in Secondary Schools, and (4) Pittsburg Achievement Renaissance. A second set of programs were said to be city-based: (1) Pittsburg New Futures Initiative, and (2) the Mayor's Commission on Families. Implemented programs were also state-supported: (1) Teacher Expectations and Student Achievement (TESA), and (2) Occupational Vocational programs such as (a) Second Chance, (b) Select Employment Trainee program (SET), (c) Summer Academy, (d) Northside Civic Development Project, and (e) Business and Finance Academy. An appendix included a description of Board goals and strategies for implementation.

The Denton and Davis (1988) report is notable for demonstrating how parents, community leaders, and the corporate sector can play a role in enhancing the achievement of minority students. Regrettably, the results of the interventions were not available at the time of the initial report nor has a follow-up study been published. The value of Denton and Davis (1988) to the present study lies its suggestion that recognizing the existence of the white-minority achievement gap is but the first volley in an allied attack to narrow it.

Haycock and Navarro (1988) discussed the programs of three California schools (Sweetwater High School, Claremont Middle School, and Bell Gardens Elementary School) that have begun to close the achievement gap between Latino and black students and other students. The "effective" schools were said to share six commonalties: (1) a determined principal, (2) demanding teachers, (3) a rigorous core curriculum, (4) parents as partners, (5) support for students, and (6) teamwork. Arguing for why the achievement gap must be narrowed, the authors cited demographic data to show that,

over the last 10 years, California's society has shifted from a predominantly white to a predominantly ethnic society. Haycock and Navarro also indicated that the number of the state's people living in poverty has increased and that most of these people are concentrated in big cities.

As a report of qualitative field research, Haycock and Navarro's (1988) portrayals of the three schools is sketchy at best. Perhaps additional information (e.g., evidence of prolonged engagement, persistent observation, and triangulation, etc.) that would better establish the study's credibility is available; it was not, however, part of the narrative. Nevertheless, Haycock and Navarro (1988) contribute to the present research by suggesting that schools effective in narrowing the white-minority achievement gap attain such results in systemic, identifiable, and replicable ways.

In a report from the Bay Shore (NY) Union Free School District, Glassman and Roelle (1990) found that efforts specifically intended to close the achievement gap between white majority students and Hispanic and black minority students had positive effects. Initiated as a total district improvement effort, the process used focus groups, survey data, community input, and school-based feedback. Programmatic changes included implementation of full-day kindergarten and new mathematics and reading curricula. Glassman and Roelle reported that black students exhibited higher achievement when they experienced the programmatic changes. Mathematic achievement scores, as measured by the New York State tests, ranged from an average of 4 to 10 normal curve equivalent (NCEs) higher among blacks who experienced the new curriculum. Similarly, test scores of blacks who experienced the new reading curriculum were significantly higher than those who did not experience the new program. The authors said:

In the cohort that had not experienced new curricula and full-day kindergarten a rather low 58.1 percent of the students passed the reading of the PEP, whereas

in the cohort that did have the new programs that percentage jumped to 78.4 percent...The achievement data for Hispanic students essentially reaffirmed everything previously stated with respect to black student achievement (pp. 9-10).

Regrettably, Glassman and Roelle (1990) did not elaborate on the new curricula. Full-day kindergarten, on the other hand, is self-explanatory. This study is useful in demonstrating one district's commitment to and subsequent success in raising the academic achievement of minority students. Full-day kindergarten and mathematics and reading curricula that promote learning for all are seen as key interventions.

Jennings (1992), in Closing the Achievement Gap: A Model for Success, reviewed three early intervention programs with demonstrated effectiveness in improving the academic achievement of disadvantaged minority students. A synopsis of each program and the results of its implementation in a school district of predominantly black students follows: The Responsive Early Childhood Education Program (RECEP) was designed "to build upon gains that low-income children had made in Head Start or similar preschool programs. . . by: (1) increasing their learning of the basic skills of language and mathematics; (2) developing their problem-solving abilities; and (3) fostering positive attitudes towards learning (Goldsboro, 1987, p. 2,6)." Implemented in the Goldsboro (NC) City Schools, the program served a single cohort of all K-3 students in the district through (a) instructional services, (b) health services, (c) social services, (d) nutritional services, (e) psychological services, and (f) parent involvement. After four years, the results of the intervention, obtained through a comparative analysis of matched sample CAT student score data, indicated that the RECEP students scored better than or equal to 49.1% of all other comparable students in reading, better than or equal to 54.1% of all others in language arts, and 53.1% in mathematics. Moreover, RECEP students scored 3.4 NCE scores higher than

the state's population in reading, 4.2 NCEs higher in language arts, and 1.6 NCEs higher in mathematics. No information was provided on the gains of both groups from first grade to third grade. Jennings concluded that RECEP contributes to the knowledge base about the design and implementation of effective early grade intervention programs by (a) emphasizing immediate and long-term academic achievement goals, (b) using a comprehensive approach to address each child's academic, physical, emotional, and social needs, and (c) enhancing the existing curriculum through emphasis on the instructional approach, not the instructional materials.

Jennings' reviewed the Columbus (OH) Public Schools' experience with the Reading Recovery program. Designed as an early intervention, Reading Recovery targets for instruction the lowest 20% of first grade student readers; however, the central component of the program is the teacher (Allington & Johnston, 1989; Pinnell, DeFord, & Lyons, 1988). Program teachers are exposed to a year of staff development during which they are taught "to observe and record children's literacy development and to tailor instruction specifically to the child's needs (Allington & Johnston, p. 343)." The results of Reading Recovery in the Columbus Public Schools in 1985-1986 indicated that 73% of the student participants were successfully discontinued from the program at various times during the school year and received no further treatment. Jennings stated that student participants performed better than comparison groups (remedial reading and random sample) in all seven program measures. A "sustained effects" study showed similar success (p. 8). Jennings suggested that among the qualities of the Reading Recovery program that make it an effective early intervention are its (a) emphasis on accelerated learning, (b) investments in the teacher rather than in materials, and (c) establishing high expectations through requiring mastery of skills to exit the program.



Jennings described the Success for All program as an early intervention based on prevention and immediate, intensive correctives. Quoting program designers, Jennings said that Success for All offered "a comprehensive approach emphasizing early education, improvement in instruction and curriculum, and intensive intervention at the earliest possible stage when deficits first begin to appear (Madden, Slavin, Karweit, Dolan, & Wasik, 1991, p. 2)." Program components included (a) reading tutors, (b) reading programs, (c) eight-week reading assessments, (d) preschool and kindergarten, (e) family support team, (f) program facilitator, (g) teachers and teacher training, (h) special education, and (i) advisory committee. Jennings reported that the results of a multi-year (K-3) evaluation utilizing student groups matched on the basis of Boehms and Metropolitan and California Achievement Test (NRT) scores indicated that "Success for All students achieved positive outcomes in most cases (p. 9)." The author suggested that a major strength of the program was its comprehensive, sequential nature and its corroboration of the Peterson (1987) findings.

Jennings concluded that a synthesis of the research on early grade intervention tends to suggest the following: (a) early learning serves as a foundation for subsequent learning, (b) early grade intervention involves the relationship between the learning environment and the child's cognitive development, (c) early intervention can make a significant difference in the developmental status of young children faster than later remedial efforts, (d) there appears to be a difference between early grade intervention and remediation, (e) early grade intervention includes involvement and empowerment of parents and teachers, (f) belief in the conceptual aspects of early grade intervention listed promises to create radical changes in the organizational structure of urban elementary schools, (g) short-term and long-term achievement goals are important, (h) successful programs place more emphasis on instructional approaches than

instructional materials, (i) successful programs utilize ongoing staff development, (j) instructional leadership extends beyond the school principal, (k) monitoring is ongoing, and (l) successful interventionists plan collaboratively (p. 10-12).

Despite the omission of important facts in his reviews of the research (e.g., the site of the Success for All program evaluated), Jennings (1992) may be credited with articulating ostensibly parallel aspects in the conceptualization, design, and implementation of early grade intervention. Most significantly, the initial success of the programs reviewed by Jennings is shown to persist over time. Thus as a means of closing the white-minority achievement gap and keeping it closed, Jennings' findings are particularly important to the present research. In a paper presented at the annual meeting of the American Educational Research Association, Ramey (1992) identified and analyzed classroom practices that correlated with measures of ethnic achievement gap reduction. Earlier research (Ramey, Hillman, & Matthews, 1982; Ramey, 1984) suggested that variables obtained from classroom observation (student activities and teacher activities) and responses to a teacher questionnaire were predictive of school achievement. Moreover, she demonstrated that a measure of ethnic achievement gap reduction could be obtained from regressing minority and majority student NRT score differences from one year over those of the preceding year. Then, by saving the residual scores and dividing each by its standard error of estimate, a "gap reduction index" could be derived (Ramey, 1992, p. 3). Using Linear Structural Relations analysis (LISREL VI), Ramey modeled the relationship between the predictor (classroom observation and teacher questionnaire response) variables and outcomes, degree of gap reduction in reading, math, and language computed from scores on the fall 1981 and spring 1982 administrations of the CAT. Her sample consisted of 47 intermediate (grades four through six) classrooms and 36 primary (grades two and three) classrooms.

For the intermediate grades, Ramey reported that the following variables appeared most predictive of reading gap reduction: (1) percent of teacher time spent one-to-one with students, (2) percent teacher time spent in organization activities, (3) percent teacher time spent in interactive instruction, and (4) average percent of student time spent in activities related to subject being taught. Together, these variables accounted for 45.1% of the variance in the gap reduction index for intermediate reading achievement. For intermediate math, the following variables, accounting for 47.4% of the variance, appeared most predictive of math gap reduction: (1) average percent of student time spent in seatwork, (2) percent teacher time spent interacting with whole class, (3) percent teacher time spent on discipline, and (4) teachers' perception of usefulness of district curriculum (negative relationship). Accounting for 46.4% of the variance, the following variables were reported to be most predictive of intermediate language gap reduction: (1) percent teacher time spent interacting with whole class, (2) percent teacher time spent in organization activities, (3) average percent of student time spent being tested (negative relationship), and (4) teachers' perception of usefulness of district curriculum (negative relationship).

For the primary grades, Ramey reported that the variables most predictive of reading gap and math gap reduction were: (1) average percent of student time spent in seatwork, (2) average percent of student time spent in "off task" (activities not related to subject being taught), (3) percent teacher time spent in lecture (negative relationship), and (4) average percent of student time spent being tested (negative relationship). These variables accounted for 43.6% and 50.1% of the variance in the gap reduction index for reading and math achievement respectively. (Language gap indexes were not computed because grade 2 students had no language pretest.)

Commensurate with the research literature on learning contexts and instructional strategies (Garner, 1990; Good, Grouws, Mason, Slavings, & Cramer,

1990), Ramey concluded that: (1) Intermediate reading gap reduction is optimized in a classroom with considerable one-on-one interactive instruction and substantial time spent organizing to keep other children involved in reading-related activities; (2) Intermediate math gap reduction is optimized in a well-disciplined classroom in which the teacher continues to interact with students as they work at their seats on teacher-developed assignments; (3) Language gap reduction is optimized in classrooms where the teacher interacts with the entire class through structured (nontest) activities using teacher-developed materials (pp. 4,5). Relative to primary grade instruction, Ramey concluded that optimal classroom for math and reading gap reduction is "permissive enough to allow students' attention to wander off task but disciplined enough to require that they remain in their seats (p. 6)." Ramey was particularly reflective on her finding that more off-task behavior narrows the achievement gap at the primary level. Referencing Piaget (1962) and Sylva, Bruner and Genova (1976), Ramey suggested that "off-taskness" is a form of play and that play reduces frustration and fear of failure. Moreover, commensurate with Christie and Johnsen's (1983) research, Ramey suggested that providing the opportunity for disadvantaged minority children to engage in play gives them a "leg up" compared with minority children who do not have the opportunity to "make believe:" "In the case of young minority students, off-task time may be among the few times that permit such a luxury (p. 8)."

Ramey (1992) may be credited not only with identifying classroom practices that correlate with measures of achievement gap reduction but for trying to figure out why the practices work. Additionally, she may be seen as the first education researcher to devise an index of the achievement gap utilizing yearly comparisons of NRT difference scores of minority and majority students. One must also remark at the large amount of variance accounted for by the school and classroom predictor variables employed by Ramey. Noting Bobbett, French, and Achilles' (1991) observation that accounting for

more than 25% of the variance in an achievement-related outcome is a rare research event, Ramey's near doubling (in some cases) of that figure is significant. Finally, Ramey's work is commensurate with Effective Schools literature inasmuch as it employs only variables influenced by teachers and the classroom. This writer suggests that several of these variables (e.g., one-to-one setting with teacher, time spent "off-task") may be even more within teachers' ability to control when class-size is small. Thus Ramey's findings are seen as supportive of the notion that small class-size may be an effective early intervention for reducing the white-minority achievement gap.

Parsons (1993) reported on a program designed "to serve the nation's minority and economically disadvantaged students, who typically are left behind when it comes to 'gateway' courses to higher level math and rewarding and lucrative career choices (p. 6)." The program, Equity 2000, began as a pilot in the Fort Worth (TX) School District and is now in place in five additional U. S. sites, including Prince Georges County (MD), Milwaukee (WI), Nashville (TN), Providence (RI), and San Jose (CA). Building from Palavin's (1988) research showing that "differences in college attendance rates of white and minority students were virtually eliminated among students who took advanced mathematics courses," the program is said to involve schools, communities and higher education in a systemic effort to assist students in successfully completing high-level math courses (p. 6). Among the implemented strategies were: (1) Algebridge, a teaching tool developed by the Educational Testing Services to facilitate the transition from the concrete operations of arithmetic to the abstract concepts of algebra; (2) Summer academies designed to (a) introduce students to upcoming lessons to assuage anxiety experienced before entering school in the fall and (b) teach parents how to support their children through "family math" activities; (3) Cooperative learning; (4) Staff development for teachers designed to (a) change their beliefs about minority and disadvantaged students through Multicultural

Institutes and (b) train them to better communicate their subject matter to diverse groups of students. Parsons quoted Vinetta Jones, national director for Equity 2000: "It's important that people believe kids can do it. I feel this has the single largest impact on whether kids succeed or not (p. 6)." The only reported result was derived from a survey of "more than 500" middle and high school teachers who participated in Equity 2000 institutes designed to explore and change their beliefs about who can do precollegiate mathematics. Parsons indicated that, before the institutes, teachers, on average, believed that 36% of their students were capable of passing algebra and that 32% of their students could pass geometry; after the institutes, the percentages climbed to 64% and 59%, respectively (p. 6).

Although there is currently no evidence to suggest that minority students exposed to Equity 2000 are achieving success in high level mathematics courses, to the extent that teachers' expectations are predictive of student achievement, Parsons' (1993) report is significant. In involving teachers, parents, and higher education, Equity 2000 is clearly a collaborative effort. Unlike programs reported by Sang (1987), Murphy (1988), Glassman and Roelle (1990), and Jennings (1992), however, the focus of Equity 2000 is exclusively on secondary school students. Whether or not the program effectively reduces the white-minority achievement gap in terms of successfully completing precollegiate mathematics courses and ultimately, collegiate mathematics courses, awaits disclosure.

#### Summary of the Attempts to Reduce the White-Minority Achievement Gap

In summary, the literature reviewed in the previous section coheres around the Effective Schools concepts that all children can learn and that schools are responsible for that learning. As such, these reports represent the far end of the explanatory continuum which began with home and family conditions to explain the white-minority achievement gap (Cultural Deprivation Theory), moving to the interaction of the home

and school (Cultural-Difference Theory), and finally to schools themselves as the locus of responsibility (Effective Schools Theory). When Cultural Deprivation is viewed as the problem, evidence suggests that the integration of culturally deficient minority students with more advantaged white students may be an effective long-range strategy; however, barriers to implementation (e.g., public acceptance of forced busing) remain. Although Cultural-Difference theorists recognize the impediments of deficient (defined as different from the mainstream culture) home conditions for minority achievement, they also argue that an understanding sufficient to reduce the white-minority achievement gap must take into account what happens when home interacts with school. For example, when English is not the primary language of the home, narrowing the achievement gap may require that the affected student be taught English first. Effective Schools proponents focus on the school practices (correlates) that appear to raise the achievement of all students. Indeed, a school in which student success may be predicted based on ethnicity, gender, socioeconomic status, or any other categorical variable is not, by definition, an effective school.

What then can be said of the reports written from the Effective Schools perspective reviewed in the previous section? Except for the Ramey (1992) study, included for its theoretical and methodological contribution to the present research, each report describes programmatic interventions in specific schools and school districts in the U. S. The commonalities of these interventions are remarkable. All or most of the programs reviewed: (1) Acknowledge publicly the existence of a white-minority achievement gap as manifest by (a) NRT score disparities between whites and minorities increasing proportionally with grade-level or (b) precollegiate mathematics course selection favoring white students; (2) Recognize the need for pervasive intervention and district-level leadership; (3) Recognize the primacy of people and training over materials as manifest by (a) collaborative efforts involving

teachers, parents, and often community institutions and institutions of higher education, (b) staff development and institutionalized support for interventionists, and (c) student lifestyle change efforts supported by community resources; (4) Recognize the importance of establishing high expectations for student achievement as manifest by (a) mastery learning or similar programs, and (b) implementation of TESA training or similar programs for teachers; (5) Focus on the curriculum as manifest by (a) reductionist curricula or alternately (b) accelerated curricula with expanded opportunities; (6) Recognize the need to address jointly skills and attitudes of both interventionists and students. Three of the nine studies (Glassman & Roelle, 1990; Jennings, 1992; Ramey, 1992) recognize and report on the relative efficacy of early (primary school) intervention. Two of these (Jennings and Ramey) suggest that small class-size may be a significant factor in reducing the white-minority achievement gap. Accordingly, the next section focuses on studies that support this hypothesis.

#### Does Small Class-Size Reduce the White-Minority Achievement Gap?

Three related studies form the foundation from which the present research proceeds: (1) Tennessee's Student Teacher Achievement Ratio (STAR) Project (Word et al., 1990), (2) a Lasting Benefits Study (LBS)(Nye et al., 1992a; 1992b), and Project CHALLENGE (Achilles et al., 1992). The extent of their contribution to an understanding of small class-size as an early intervention strategy for the reduction of the white-minority achievement gap is the focus of this section. Following a discussion of the political context of Project STAR, this writer will describe the sample selection, study design and data collection, and findings pertinent to the present research.

#### The Political Context of Project STAR

Folger (1989/1992) explained that Project STAR had its genesis in the 1983 controversy surrounding Tennessee Governor Alexander's Better Schools Program. The Tennessee Education Association (TEA) opposed the Career Ladder program component,



which would reward teachers commensurate with student outcomes. The TEA argued that students would be better served by lowering class size in the early elementary grades from the existing 25 to 21. Although TEA's proposal was ultimately rejected by the Governor and the Legislature, House Representative Steve Cobb successfully sponsored a bill in 1985 providing for a "demonstration" to study the effects on student achievement and development of a decrease to 15 students classes in Grades K-3.

Specified by the legislation were a number of key conditions: (a) There would be a large sample of schools representative of various locations (i.e., urban, rural, suburban, and inner city); (b) Schools would be drawn from all parts of the state; (c) The study would be longitudinal and follow a single cohort of students over a 4-year period from kindergarten through the third grade; (d) Teacher training in how to maximize instruction in small classes would be provided; and (e) The project would be funded at three million dollars a year with the state picking up the tab for the extra teachers and aides that would be needed. STAR was to be "a definitive study which would establish for Tennessee (and for other states with similar early elementary school programs) the size of the class size effect (Folger, p. 9)."

In May 1985 the legislation passed. Under the leadership of the State Department of Education, four major universities--Memphis State University, Tennessee State University, University of Tennessee, Knoxville, and Vanderbilt University--were contracted to design, collect and analyze the data, and develop of the final report of the study. Dr. Roy Forbes, Dr. Jeremy Finn, and Dr. Doris Ryan served as the project's external advisory committee. Dr. Elizabeth Word, an experienced elementary principal, was employed by the State Department of Education to manage the project. Oversight was shared among the SEA, the university researchers, a representative from the State Board of Education and one from the State

Superintendents' Association. The implementation of the project began in August, 1985. First, however, the sampling issue had to be addressed.

#### Selection of the Project STAR Sample

Given the legislative requirement for equal geographic representation, the State Commissioner of Education invited all systems in the state to participate. About one-third of the districts, representing 180 schools, volunteered. Of these, 100 schools met the criteria for inclusion in the study. The within-schools design called for schools sufficiently large to permit random assignment of teachers and students to one of the two treatment groups, small (13-17 pupils) or regular with a full-time teacher aide (21-25 pupils) or to a control group, a regular (21-25 pupils) class. Seventy-nine schools were selected with 328 kindergarten classes: 128 small classes (S), 101 regular (control) classes (R), and 99 regular-with-aide (RA) classes. The final sample was reduced to 76 elementary schools when three schools requested to be dropped. At its maximum, STAR contained about 7,000 pupils. Post hoc comparison of the participating school districts indicated that they were larger than the state average, had moderately higher per-pupil expenditures and teacher salaries, but were nearly identical in teacher-pupil ratios and percentages of teachers with degrees beyond the bachelors. Participating schools scored only slightly below the state average in second-grade reading and mathematic standardized test scores. "The project staff concluded that the sample schools in Project STAR were representative of all schools in the state and that results from the project could be generalized to all schools in the state (Folger, p. 10)."

#### Project STAR Study Design and Data Collection

To minimize the effects of variations in school conditions across the state (e.g., socioeconomic status, community resources, availability of instructional materials, school administration, policies, per-pupil expenditures), a within-school design was

chosen. At least three classes, one for each of the two experimental conditions and one control class, were housed within each participating school. Small classes averaged 15 pupils (17 maximum) while R and RA classes averaged 23 pupils (25 maximum). Accordingly, the class was to be the unit of analysis. Each year project staff randomly assigned teachers to one of the three class types. Students were randomly assigned to class type in their kindergarten year and stayed with that type for the remaining three years. By the fourth year of the project, about one-third of the students had been in the same class type all four years, the remaining two-thirds were replacements and additions who were replaced via a randomization process.

The state-required Stanford Achievement Test (SAT) was the *norm-referenced* test (NRT) instrument utilized during the study. Additionally, *criterion-referenced* tests (CRT) in reading and math (Basic Skills First) based on the learning objectives in the Grades 1-3 curriculum were employed. Using both the norm referenced (SAT) and the criterion referenced Basic Skills First (BSF), students were tested in April of each year. Standardization of test conditions was assured through the use of trained substitute teachers as proctors. Fewer than one-third of one percent of the testing periods failed to meet the minimum testing criterion. Although the primary focus of the study was on effects on student achievement and development (through such things as attendance), researchers collected abundant information on classroom activities and on teachers' opinions and perceptions of the different classroom conditions as well. Similarly, project staff collected teacher, administrator, school, and class demographic indicators and other measures of student behavior (e.g., the Self-Concept and Academic Motivation Inventory (SCAMIN), attendance records, and grade retention information).

#### Validity Issues and Project STAR Findings

Project STAR researchers controlled for threats to validity through randomization of teacher and student assignment to class type within each school.

Although school officials were responsible for assignment, placement was monitored by project personnel. Moreover, new entrants were placed according to a random assignment plan provided by the researchers. High SES students tended to be slightly over-represented among new entrants to small classes, but researchers determined the effect of those higher proportions of high SES students in small classes could have increased the difference in small and regular achievement scores by no more than five percent. Additionally, students who were "pulled out" of both small and regular classes for special instruction posed another threat to validity. When this occurred, the average size of affected small classes fell to 13.9 while that of regular classes fell to 19.7. Nonetheless, the researchers observed that the ratio of small to regular classes was consistent with project guidelines. "In general, these threats to validity do not appear to be large enough to affect the basic conclusions of the study, or to invalidate any of the major findings, with the possible exception of the finding of small and largely nonsignificant differences in student achievement between aide and regular (control) classes (Folger, p. 13)." Accordingly, if pupils in small classes were shown to score generally higher on cognitive outcome measures than pupils in regular classes, their higher scores may be attributed to small class participation.

#### Project STAR Findings Pertinent to the Present Study

Finn and Achilles (1990) calculated the means of the outcome measures for each class and then separately for white and minority students in each classroom. Multivariate analysis of variance of the entire first-grade sample (cross-section analysis) and repeated measures analysis of variance on the subset of pupils who participated for both kindergarten and first grade (longitudinal analysis) were conducted. For the cross-sectional analysis, the statistical design treated schools as a random factor, nested within the four locations (inner city, urban, suburban, and rural) and crossed with class type (small, regular, and aide) and race (white or

minority). Mean scores for whites and minorities for each class were employed as the basic units of analysis.

First-grade cross-sectional results indicated that minorities scored significantly below whites on all cognitive measures. However, students of both races enjoyed higher scores on every measure if they belonged to a small-size class as contrasted with a regular-size class. The effect sizes (ES) were about .25 among students and ranged from about .33 to .66 among class means. For some outcomes, however, significant interactions of type with race indicated that reduced class size did not have the same impact on white and on minority students. Specifically, the advantage of being in a small class was greater for minority students than for whites on all measures. For example, white students in small classes enjoyed a 8.6 point (.15 sigma) advantage on the SAT reading scale; minority students in small classes outdistanced their peers by 16.7 points (.35 sigma), more than doubling the effect size. Minority students in small classes passed the BSF reading test at a rate comparable to their white peers. Suburban minorities fared particularly well: The average percentage of minority students in small classes who pass the BSF mathematics test was 90.3. This was well above the pass rate for suburban minorities in other class types (70.7 percent), and slightly above the average of whites passing any type of class in any location.

Longitudinal analysis of data on students participating both in kindergarten and first-grade was conducted. Due to some students' nonattendance in kindergarten and to exchanges between regular and teacher-aide classes, the longitudinal sample consisted of only about 35 percent of the cross-sectional sample. The repeated measures analysis of variance (ANOVA) yielded results consistent with the cross-sectional findings. Differences among the four geographic locations and between whites and minorities were statistically significant on all measures. Small classes showed a decided advantage

over the other class types on all scales in both grades. In terms of class size and one-year gains, small classes maintained or increased their superiority over regular classes, especially for minorities. While white students' gains in mathematics were about the same for each class type, one-year minority gains in mathematics were notably larger in small and teacher-aide classrooms. In reading, both white and minority students gained more from small class participation but the minority student gains were comparatively greater. Finn and Achilles concluded:

These outcomes may be viewed in terms of a white-minority achievement gap. In kindergarten, the race difference on word study skills is 11.0 points in small classes, 10.8 in regular classes, and 8.6 in teacher-aide classes--all about the same order of magnitude. By the end of first grade, the white-minority differential was 28.8 points for students in small classes and 41.0 and 49.4 points in regular classes, respectively. The 41-point difference is closest to that which might have been expected if there were no intervention. However, with students attending small classes for 2 years, the difference was reduced to 28.8 points. This did not occur at the sacrifice of performance among whites, who also gained more in small than regular classes, but by giving an extraordinary boost to minority pupils. The same pattern is seen for reading and, to a lesser extent, for mathematics as well. It must be emphasized that these results are not confirmed by statistical tests. However, classes of 15 pupils appear to have benefited both white and minority students in terms of 1-year gains in reading and mathematics, with the greater relative benefits accruing to minorities (p. 573).

STAR researchers, however, did not continue to analyze white/nonwhite differences in detail, and specifically did not carry out extensive analyses specifically using CRT results. The STAR use of classroom as the unit of analysis may have understated within-class variances and also there was the practical problem of "average" gains with only 1-3 minority pupils in a classroom.

#### From Experiment (Project STAR) to Field Research (LBS)

Project STAR has been called "the most significant educational research done in the US during the past 25 years (Orlich, 1991, p. 632)." Similarly, STAR researchers have made unconditionally positive statements about the project: "This

experiment yields unambiguous evidence of a significant class size effect, at least in the primary years" (Finn et al., 1990, p. 135) and "This research leaves no doubt that small classes have an advantage over larger classes in reading and mathematics in the early primary grades" (Finn & Achilles, 1990, p. 573). Nevertheless, without follow-up studies the extent to which student gains attributed to small-size classes in the primary grades persist after they return to regular-size classrooms remained a matter of conjecture. Tracking student achievement beyond third grade became then a primary focus of the Lasting Benefits Study (LBS). Essentially, the large-scale experiment of STAR (1985-1989) became the field research of LBS (1989-1993) and perhaps longer.

In a paper presented at the 1992 conference of the American Educational Research Association (AERA), Five years of small class research: Student benefits derived from reduced student/teacher ratios, Nye et al. (1992a) report on STAR students one year removed (fourth grade) from the small class-size treatment. In a second paper submitted for the Mid-South Educational Research Association (MSERA) Outstanding Paper Competition, Small is far better, Nye et al. (1992b) report on STAR participants two years after (fifth grade) the original experiment. After discussing a design problem inherent in the LBS longitudinal research, this writer will review the sample and design, data collection and analysis procedures, and findings pertinent to the present study for each LBS report.

#### Problems with LBS Longitudinal Research

Nye et al. (1992a) acknowledged a relative design weakness of LBS vis-a-vis Project STAR. Once students from small classes returned as fourth graders to regular-size classrooms, they were indiscriminately grouped with non-small class-size students. Controlling for "teacher effects," as in the STAR experimental design which provided for the class (average) as the unit of measure, thus became problematic. LBS

researchers attempted to address the issue by using groups-within-class mean scores as the unit of measure. In other words, only former STAR students were grouped within each fourth grade class and according to their previous STAR class type. Nye et al. (1992a) admitted:

This use of groups-within-class mean scores will become more tenuous as students progress through the system considering the percent of former STAR students in any given class probably will diminish, and the number of classrooms containing STAR students will increase. . . Tracking the student sample, data collection and analysis continue to be a major longitudinal concern (p. 4).

While presumably doing little to placate the critics of educational research who demand a high degree of causal certainty, the complex nature of longitudinal studies makes them extraordinarily vulnerable to uncontrolled factors that threaten their validity, both internal and external (Keeves, 1988, p. 120). On this note, Kratochwill (1978, p. 11) observed that the quest for both high internal and external validity can operate at odds with each other. The requirement to achieve a high level of internal validity can demand the exercising of tight experimental controls on the collection of data that so distort the natural setting in which development occurs that the external validity of the investigation is threatened. (Presuming a desire to track achievement beyond the small-class intervention years, one way that Project STAR researchers could have controlled for threats to validity associated with age, time of measurement, and cohort would have been to matriculate more than one cohort of students through the four years of intervention. Obviously, associated costs would have been commensurately higher and possibly altogether prohibitive.) Threats to validity notwithstanding, Nye et al. (1992b) asserted that "the LBS results are informative and an important contribution to the analysis of class-size intervention and public policy decision making (p. 7)."



#### Fourth Grade LBS Sample and Design

The sample for the fourth grade analysis included 4230 students, of whom 1412 were from Grade 3 Project STAR small (S) classes, 1250 from regular (R) classes, and 1568 from regular-with-aide (RA) classes. A sizable proportion of minority students excluded due to testing changes reduced minority representation to 20.2 percent of the total. Commensurate with STAR, students were designated as either Inner city (n=326), Urban (n=257), Suburban (n=1055), or Rural (n=2592). The research question posed was: "Do students who received the benefits of being in a small class condition (as shown in STAR), and who were in a small class at least in grade three (1988-89) retain any of those benefits after being in a "regular" (1:25 or so) class for one year (grade 4), for two years (grade 5), and for three years (grade 6) (p. 4)?"

#### Fourth Grade LBS Data Collection and Analysis

Nye et al. (1992a) reported that during the first year of LBS, Tennessee discontinued use of the SAT and BSF achievement tests in lieu of the Tennessee Comprehensive Assessment Program (TCAP). The TCAP battery included both norm-referenced test (NRT) component and a criterion-referenced test (CRT) component. The NRT component measured students' abilities in reading, language, study skills, science, and social science. Designed to correspond to the Tennessee curriculum, the CRT component measured students' mastery levels (i.e., mastery, partial mastery, or non-mastery) regarding the language arts and mathematics content domains.

Through multivariate analysis of variance (MANOVA) scaled score means for the three STAR class types were compared. The analysis examined mean score differences among the class types, mean differences among the four school locations, and the interaction between class types and locations. The MANOVA model controlled for effects

of school location (fixed) and classes (random). Student scores from three achievement subsets for the LBS were compared separately. Two of the subsets included scores from both the NRT and the CRT components of the TCAP: Set 1 was Total Reading and Total Language (NRT scores) and number of domains mastered in Language Arts (CRT). Set 2 was Total Math, Total Science (NRT scores), and number of domains mastered in Mathematics (CRT). Set 3 was Study Skills and Social Science (NRT scores). Post-hoc analyses were employed for two particular contrasts: (1) Small class vs. Regular class and (2) Regular-with-aide vs. Regular class.

#### Fourth Grade LBS Findings Pertinent to the Present Study

The significant question is: Did the reduction of the white-minority achievement gap attributed to small class-size demonstrated after two years in Project STAR persist into grade four? Despite researchers' inability to make statistical comparisons between STAR K-3 SAT scores and LBS Grade 4 TCAP scores, the cross-sectional analysis of fourth grade student data tended to corroborate the STAR findings with regard to race, small classes, and achievement. When an approximate analysis to identify any interaction of race with class type on the achievement measures was conducted, Nye et al. found no statistically significant differences. The researchers observed that:

Although Project STAR results on K-3 test scores indicated that the most significant S-class advantages were found with minority students, the LBS fourth-grade results showed that the positive effects of small classes were equally significant for minority and white students, in all school locations. . . [T]he S-class students always outperformed students from R and RA class types. There is a consistent difference between white and minority students that favors white students, but the S class is consistently higher for both groups (p. 9).

A single example illustrates this conclusion: The mean Total Reading score of white students one year removed from small classes was 714.0. For minority students from small classes, the mean was 685.9, as contrasted with the 676.1 mean of minority

students from regular classes. Although the 28.1 difference between (S) whites and (S) minorities was large, it was nearly 10 points smaller than the gap between small class white and regular class minority students. A similar pattern was displayed across every measure of achievement.

In conclusion, the results from the fourth-grade LBS sample indicated that (1) the white-minority achievement gap may be reduced; (2) the reduction in the achievement gap may be attributed to small class size in the third grade; and (3) more longitudinal research is needed. Accordingly, achievement test data from Grade 5 were collected and analyzed.

#### Fifth Grade LBS Sample and Design

As reported by Nye et al. (1992b), the sample for the fifth-grade analysis included 4649 students, of whom 1578 were from Grade 3 Project STAR small (S) classes, 1467 from regular (R) classes, and 1604 from regular-with-aide (RA) classes. Due to some grade 5 pupils having entered middle schools, Nye et al. reported that analyzing for location "no longer seemed feasible (p. 8)."

#### Fifth Grade LBS Data Collection and Analysis

The TCAP battery continued to be the instrument by which student achievement was measured. Scaled score means for each of the three TCAP subsets employing class type as a fixed effect were compared through analysis of variance (ANOVA). The analysis examined mean score differences among the three class types (S), (R), and (RA). Location (see above) was not a main effect for the fifth grade analysis.

#### Fifth Grade LBS Findings Pertinent to the Present Study

There was a statistically significant ( $p \leq .01$ ) class type difference across all three TCAP subsets. An examination of the scaled score differences between (S) and (R) classes, and between the (RA) and (R) classes, indicated that the students who had attended small STAR classes had an achievement advantage over the regular-size

classes. Nye et al. reported that "The positive effects from involvement in a small-size class still remain pervasive two full years after students returned to regular-size class (p. 8)."

#### Sixth Grade LBS Findings Pertinent to the Present Study

The sixth grade LBS sample/design and data collection procedures remained consistent with the fifth-grade study. Although inferential statistical analyses on the STAR students who have completed the sixth-grade are not available at this time (November, 1993), descriptive statistics suggested a diminished effect of (S) on minority advantagement, where advantagement is defined as comparatively smaller white-minority gap scores in (S) than in (R) or (RA). In fact, the grade six gap scores on TCAP mean number of domains mastered in mathematics and language arts (CRT measures) were consistently lowest for the (RA) samples. Similarly, on all but two of the six TCAP NRT measures (Total Language and Total Mathematics), gap scores were lowest for the (RA) students. Effect sizes, however, were uniformly larger for (S) vs. (R) contrasts than for (RA) vs. (R) contrasts suggesting the greater benefit of (S)/(RA) to (R) as a "treatment."

#### From Field Research (LBS) to Policy Implementation: Project CHALLENGE

Beginning in 1989-1990, Tennessee legislators allocated funds to LEA's in some of the poorer counties of the state for the purpose of improving student performance. One strategy--Project CHALLENGE--was to reduce the class size in 17 districts in grade K-3 to approximately 1:15. Project CHALLENGE thus put into practice the results of the statewide STAR experiment.

Lacking formal design (randomization, testing, etc.), CHALLENGE was not itself an "experiment," rather the study was a post-hoc review and analysis of grouped (e.g., school system) data. There was no way to attribute gains or losses to class-size reduction if contemporaneous interventions were occurring. Other threats to validity

(e.g., student ability, excellent teaching, test variation) were also unaccounted for. The researchers suggest that "Only with several years of results can a trend become evident (Achilles et al., 1992, p. 12)."

Using grade two data, CHALLENGE evaluators developed the average rank of the 17 systems on the average system score on the TCAP. Of the state's 139 systems, the average rank for the CHALLENGE systems on Reading was 99; in Math it was 85. In 1990, these ranks were based on students' one-year participation in a small class. In 1991, ranks were developed based on two years of participation (grades 1 and 2). Over the next two years, K and grade 3 were added until CHALLENGE districts included small classes for four years of participation.

After two years of (S) treatment (1991) there was minimal change in Math (85-79), but by 1992 the 17 counties had moved to an average rank of 60, slightly above the state average. Researchers conjectured that the large move in the average rank between 1991 and 1992 could have been because the students had three years of (S) treatment. On the other hand, a major STAR finding was that the largest (S) gains occurred between K and grade 1, however 1992 was the first year for K in the Project CHALLENGE--the majority of the students tested were those who had matriculated from grade 1 to grade 2. Researchers suggested that a second explanation that fits both the findings of STAR and CHALLENGE was that (S) treatment is not a remediation but than it is a prevention (Achilles et al., 1993, p. 4). The degree to which the variables of duration and incipience of (S) treatment contribute to either a prevention or remediation hypothesis, however, remains a question that has yet to be explored in the context of a true experimental database.

### CHAPTER III

### METHODOLOGY

Commensurate with decision-oriented educational research and the use of existing databases (Cooley & Bickle, 1986; Stakenas, 1989), the present study is a re-analysis of data collected in the Student Teacher Achievement Ratio (STAR) Project (Word et al., 1990). Specifically, this study focuses on the achievement patterns of white and minority students in small, regular, and regular-with-aide class-sizes over the four years of STAR intervention. Accordingly, the purpose of this chapter is to describe the (1) STAR subjects, (2) STAR instruments and procedures, and (3) procedures employed in the present re-analysis.

#### STAR Subjects

The Project STAR design provided for four years of randomly assigned small (S), regular (R), and regular-with-aide (RA) classroom "treatments" for a single cohort of approximately 7,000 student subjects who began Tennessee public school kindergarten in 1985 and who completed third grade in 1989. (A complete review of the STAR design is found in Chapter II.) Ideally, STAR researchers intended for students assigned to one of the three class types as kindergartners to persist in that class type all four years; however, as new students moved into the attendance zone of an experimental school, they, too, were randomly assigned to one of the three conditions. Similarly, STAR researchers encountered subject attrition through the moving away of students from experimental schools and classrooms. Consequently, not all STAR students received four years of treatment. By the final year of the project, about one-third of the students had been in the same class type all four years; the remaining two-

thirds were replacements and additions. Specific numbers of STAR subjects employed in the present re-analysis will be tabulated in a later section.

#### STAR Instruments and Procedures

Although STAR researchers collected data using instruments specifically selected for Project STAR (e.g., the Self-Concept and Motivation Inventory), the primary student achievement outcome measures were obtained from instruments mandated by the TN Department of Education Testing Program. Both norm-referenced tests (NRT's) and criterion-referenced tests (CRT's) were included. The NRT's were nationally-normed, published instruments, while the CRT's were developed by the TN Department of Education. Commensurate with the state testing program, the instruments and testing procedures employed were a function of whether the sample was in kindergarten or in grades 1- 3.

#### Kindergarten

In the kindergarten sample, only a NRT, the Stanford Early School Achievement Test (SESAT) Form 2 (The Psychological Corporation, 1985), was administered. Developed as a downward extension of the Stanford Achievement Test (SAT) Series, the SESAT is a group test intended to measure school achievement in grades K.5-1.5. Scores were obtained in (1) Sounds and Letters, (2) Word Reading, (3) Total Reading, (4) Mathematics, (5) Listening to Words and Stories, (6) Total for Basic Battery, (7) Environment, and (8) Total for Complete Battery. Although validity coefficients (Person Product-Moment) between the SESAT and the Otis-Lennon School Ability Test (.81 for the complete battery) and reliability coefficients (Kuder-Richardson 20) between the SESAT and the SAT Primary 1 instrument (.45 to .52 on specific subscales) are modest, Ackerman (1989) suggests, "At a level of aggregation that considers intact classes as a whole in comparison to state and national norms, the test may provide useful information (p. 866)." For the purposes of this study, only the

Total Reading and Mathematics scores were analyzed. According to the Technical Data Report (The Psychological Corporation, 1985) compiled by the test publisher, KR-20 r's were .93 and .81 for the total reading and total math respectively (p. 32). (See Chapter IV.) The SESAT was administered during the spring of the kindergarten year under controlled testing conditions including the use of trained substitute teachers as proctors.

### Grades 1-3

Beginning in grade 1, the Stanford Achievement Test (SAT) Primary 1 was employed as the NRT instrument. The SAT is a nationally normed group test intended to measure "the important learning outcomes of the school curriculum (The Psychological Corporation, 1985)." Scores for the Primary 1 form (administered in grades 1 and 2) were obtained in (1) Word Study Skills, (2) Word Reading, (3) Reading Comprehension, (4) Total Reading, (5) Concepts of Number, (6) Mathematics Computation, (7) Mathematics Applications, (8) Total Mathematics, (9) Language, (10) Spelling, (11) Environment, (12) Listening, (13) Basic Battery, and (14) Complete Battery. Primary 2 form, administered in grade 3, is identical to Primary 1 except for substituting "Reading Vocabulary" for "Word Reading." Kuder-Richardson (KR) 20 reliability coefficients have been obtained for each test and subtest for each form and level. These range from .85 to over .90. Alternative form reliability coefficients tend to be slightly lower than the corresponding KR-20 coefficients, but almost all are .80 or higher (Carpenter, 1989). The present study utilized only the Total Reading and Total Mathematics subscores as student achievement NRT outcome measures.

Beginning in grade 1, the TN Department of Education Testing Program provided for the use of state-developed curriculum-referenced (criterion-referenced) tests locally known as Basic Skills First (BSF) tests (TN Department of Education, 1987).



These tests were created from well-specified lists of objectives in reading and mathematics at each grade level, and can be scored either as the total number of items answered correctly, or as pass-fail. A student passes if he/she masters 80% of the objectives covered by the test items. The present study used the pass-fail method in the reported analyses. No reliability data are available for the BSF tests. The SAT and BSF tests were administered during the spring of the grades 1, 2, and 3 years under controlled conditions.

#### Procedures Employed in the Present Re-Analysis

To examine the achievement patterns of white versus minority students in small (S), regular (R), and regular-with-aide (RA) class-sizes as a function of time in "treatment" and beginning school year of treatment required: (1) identifying the appropriate subsample of students in each grade and class-size and for each achievement measure; (2) calculating the white and minority subsamples' mean scaled scores and standard deviations for each achievement measure by length of time in each class-size treatment at each grade; (3) tabulating the white and minority subsamples' mean scaled scores for each achievement measure by length of time in each class-size treatment at each grade; (4) tabulating a white-minority achievement gap as shown by differences in mean scaled scores of each achievement measure by grade by length of time in each class-size treatment; (5) tabulating achievement differences between minority small-class and white regular-class subsamples by grade by length of time in class; (6) calculating "effect sizes" to measure the "educational importance" of selected mean score achievement differences between subsamples; and (7) charting histograms to compare the achievement of white (R) and minority (R) and (S) class-sizes. Regarding procedure (6), Glass and Hopkins (1984, p. 54) have explained that "educational importance" or "effect size" (ES) can be mathematically represented as the quotient of the difference (gain or loss) between any two means and the standard

deviation of their combined distributions. It is expressed as a proportion of 1.00 sigma.

## CHAPTER IV

### RESULTS

Commensurate with the purpose of the present study to examine the achievement patterns of White and minority students in small, regular, and regular-with-aide class-sizes over varying lengths of time in "treatment" and for varying start-up years, the researcher tabulated student achievement outcomes for white and minority students as a function of years of participation in Project STAR for each class-size. The results are reported in response to the Research Question that generated them.

#### Research Question (1)

What patterns of minority student achievement (reading and mathematics NRT mean scaled scores and CRT average percentage passing) are associated with exposure to small (S), regular (R), and regular-with-aide (RA) class-sizes for one year (grades K, 1, 2, or 3), two years (grades K+1, 1+2, or 2+3), three years (grades K+1+2 or 1+2+3) and four years (grades K+1+2+3)? (In the tables that follow, the reader is advised that (1) the grade K Stanford Achievement Test Series (The Psychological Corporation, 1985) measure was actually the Stanford Early Scholastic Achievement Test (SESAT) as described in Chapter III of this dissertation, and (2) The TN State Department of Education did not provide Basic Skills First (BSF) achievement measures.)

Table 2 shows the results for (S) exposure. Except for the 4 years subsample, both SAT (the NRT component) mean scaled scores of minority subsamples which spent 1 year in (S) treatment exceeds that of subsamples exposed 2 years or 3 years. Similarly, the BSF (the CRT component) measures of (S) minority students are, except for the subsample that began in K and continued through grade 3, greater for each 1

Table 2

**Minority Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Small Class-Size**

---

<b>Achievement</b>						
<b>Measures</b>	<b>Length of Time in Small Class-size</b>					
	<b>Grade</b>	<b>1 year</b>	<b>2 years</b>	<b>3 years</b>	<b>4 years</b>	
<b>SAT Reading</b>	K	439.12	-----	-----	-----	
	1	510.18	495.35	-----	-----	
	2	571.51	569.18	561.37	-----	
	3	607.79	606.19	594.55	612.78	
<b>BSF Reading</b>	K	-----	-----	-----	-----	
	1	27.22	24.83	-----	-----	
	2	38.64	37.16	35.17	-----	
	3	31.17	30.71	28.03	32.11	
<b>SAT Math</b>	K	479.56	-----	-----	-----	
	1	521.49	503.40	-----	-----	
	2	568.91	566.93	556.46	-----	
	3	608.79	604.39	594.90	613.63	
<b>BSF Math</b>	K	-----	-----	-----	-----	
	1	38.96	35.86	-----	-----	
	2	51.90	51.90	48.49	-----	
	3	48.00	47.62	44.84	48.85	

---

year subsample than for corresponding multiple year subsamples. In fact, except for the subsample that spent 4 years in (S), increased time in treatment appears to be associated with lower achievement. On average, both SAT and BSF mean differences between 1 year, 2 years, and 3 years exposed subsamples were moderately small with effect sizes (ES)  $\leq .20$  sigma. At 18.73 points (.49 sigma), the point of maximal difference was the SAT Math mean scores between the 3 years and 4 years subsamples.

Table 3 shows the results for minorities exposed to (R). With one exception (BSF Math, 2 years), the same pattern of achievement attained by (S) subsamples was shown by (R) subsamples. Although displaying a wide range of differences across any one grade, the achievement for 1 year subsamples tended to be greater than that of multiple year students. At 16.88 points (.43 sigma), the point of maximal difference was again the SAT Math mean scores between the 3 years and 4 years subsamples.

Table 4 shows the results for minorities exposed to (RA). As in the (S) and (R) subsamples, 1 year achievement measures uniformly exceeded 2 years measures; however, 3 years subsample mean scores were greater than 1 year or 2 years measures while 4 years means exceeded 3 years means. After 1 year, then, there appears to be a positive linear relationship between length of time in (RA) and achievement. The average ES between the lowest SAT scores in the 2 years subsamples and the highest scores in the 4 years subsamples was about .32 sigma. Both BSF outcomes in the (RA) 3 years exposure at grade 2 subsample exceeded that of students who had been similarly exposed at grade 3 for all 4 years by about 5.00 percentage points (.67 sigma).

#### Research Question (2)

What patterns of white student achievement (reading and mathematics NRT mean scaled scores and CRT average percentage passing) are associated with exposure to

Table 3

**Minority Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular Class-Size**

---

<b>Achievement</b>					
<b>Measures</b>	<b>Grade</b>	<b>Length of Time in Regular Class-size</b>			
		<b>1 year</b>	<b>2 years</b>	<b>3 years</b>	<b>4 years</b>
<b>SAT Reading</b>	<b>K</b>	425.60	-----	-----	-----
	<b>1</b>	491.36	468.23	-----	-----
	<b>2</b>	558.82	546.73	555.95	-----
	<b>3</b>	595.81	590.18	592.23	607.89
<b>BSF Reading</b>	<b>K</b>	-----	-----	-----	-----
	<b>1</b>	24.35	20.41	-----	-----
	<b>2</b>	35.84	31.67	35.84	-----
	<b>3</b>	30.01	29.24	29.04	31.87
<b>SAT Math</b>	<b>K</b>	472.53	-----	-----	-----
	<b>1</b>	508.39	484.98	-----	-----
	<b>2</b>	559.20	542.68	558.21	-----
	<b>3</b>	600.62	599.73	593.00	609.88
<b>BSF Math</b>	<b>K</b>	-----	-----	-----	-----
	<b>1</b>	36.45	31.92	-----	-----
	<b>2</b>	49.67	46.71	48.56	-----
	<b>3</b>	46.08	45.07	44.78	48.24

---

Table 4

**Minority Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular/Aide Class-Size**

---

Achievement		Length of Time in Regular/Aide Class-size			
Measures	Grade	1 year	2 years	3 years	4 years
SAT Reading	K	428.61	-----	-----	-----
	1	493.18	475.33	-----	-----
	2	559.72	542.57	556.80	-----
	3	598.26	594.57	600.46	605.18
BSF Reading	K	-----	-----	-----	-----
	1	24.83	21.93	-----	-----
	2	35.81	30.74	36.28	-----
	3	30.00	29.84	29.97	31.16
SAT Math	K	469.64	-----	-----	-----
	1	509.61	494.58	-----	-----
	2	557.51	531.85	557.15	-----
	3	598.98	595.41	604.80	605.51
BSF Math	K	-----	-----	-----	-----
	1	36.81	33.13	-----	-----
	2	49.91	44.42	49.50	-----
	3	46.21	45.75	47.28	47.46

---

small (S), regular (R), and regular-with-aide (RA) class-sizes for one year (grades K, 1, 2, or 3), two years (grades K+1, 1+2, or 2+3), three years (grades K+1+2 or 1+2+3) and four years (grades K+1+2+3)? Table 5 shows the results of white student subsamples exposed to (S) treatment. For both SAT and BSF measures, 1 year subsamples uniformly achieved greater than 2 years and 3 years subsamples. White students who began (S) treatment in K and who continued all 4 years, however, showed the highest levels of achievement. The ES differences between the 4 years subsamples and the 1 year subsamples (about .16 sigma on the SAT measures and .08 sigma on the BSF measures) were small. The differences between the 3 years and 4 years subsamples, however, increase to .38 sigma and .30 sigma for the SAT and BSF respectively. Obvious in its singularity, the grade 2, 3 years subsample outscored the grade 3, 4 years subsample by 3.67 percentage points (.63 sigma) on the BSF reading measure.

Table 6 shows the results of white student subsamples exposed to (R). The pattern of relatively high 1 year mean scores with diminishing 2 years and 3 years scores, followed by the greatest achievement from 4 years subsamples, recurs in these data. Achievement ES differences between highest and lowest mean scores average about .43 on the SAT. Student subsamples in grade 2, regardless of length of time spent in (R), outscore student subsamples exposed 4 years by an average of more than 4.00 percentage points (.67 sigma) on BSF measures. The one exception is at grade 2, 1 year.

Table 7 shows the results of white student subsamples treated in (RA). The 1 year subsamples again demonstrate consistently higher levels of achievement than either 2 years or 3 years subsamples. Achievement measures for 4 years subsamples, however, remain uniformly highest. The average ES difference between the lowest and highest SAT mean scaled scores is about .44 sigma. BSF Reading measures at grade 2,



Table 5

White Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Small Class-Size

---

Achievement		Length of Time in Small Class-size			
Measures	Grade	1 year	2 years	3 years	4 years
SAT Reading	K	443.70	-----	-----	-----
	1	539.49	513.31	-----	-----
	2	599.95	570.18	578.14	-----
	3	627.15	619.10	619.07	633.43
BSF Reading	K	-----	-----	-----	-----
	1	27.81	24.55	-----	-----
	2	41.49	35.42	38.30	-----
	3	34.04	33.58	33.33	34.63
SAT Math	K	496.20	-----	-----	-----
	1	546.67	523.72	-----	-----
	2	595.46	576.87	575.46	-----
	3	629.25	621.32	620.35	635.12
BSF Math	K	-----	-----	-----	-----
	1	40.45	36.34	-----	-----
	2	54.15	51.45	51.15	-----
	3	53.18	52.55	51.86	54.01

---

Table 6

**White Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular Class-Size**

---

<b>Achievement</b>		<b>Length of Time in Regular Class-size</b>			
<b>Measures</b>	<b>Grade</b>	<b>1 year</b>	<b>2 years</b>	<b>3 years</b>	<b>4 years</b>
<b>SAT Reading</b>	<b>K</b>	<b>439.07</b>	-----	-----	-----
	<b>1</b>	<b>527.31</b>	<b>501.43</b>	-----	-----
	<b>2</b>	<b>591.97</b>	<b>579.68</b>	<b>578.28</b>	-----
	<b>3</b>	<b>621.59</b>	<b>618.98</b>	<b>610.61</b>	<b>627.51</b>
<b>BSF Reading</b>	<b>K</b>	-----	-----	-----	-----
	<b>1</b>	<b>26.78</b>	<b>24.16</b>	-----	-----
	<b>2</b>	<b>40.64</b>	<b>38.10</b>	<b>38.66</b>	-----
	<b>3</b>	<b>33.35</b>	<b>33.29</b>	<b>32.09</b>	<b>34.06</b>
<b>SAT Math</b>	<b>K</b>	<b>488.34</b>	-----	-----	-----
	<b>1</b>	<b>535.53</b>	<b>514.83</b>	-----	-----
	<b>2</b>	<b>588.63</b>	<b>578.50</b>	<b>572.95</b>	-----
	<b>3</b>	<b>624.26</b>	<b>623.73</b>	<b>615.92</b>	<b>630.43</b>
<b>BSF Math</b>	<b>K</b>	-----	-----	-----	-----
	<b>1</b>	<b>39.78</b>	<b>36.15</b>	-----	-----
	<b>2</b>	<b>53.23</b>	<b>50.57</b>	<b>50.75</b>	-----
	<b>3</b>	<b>52.16</b>	<b>52.26</b>	<b>50.17</b>	<b>53.29</b>

---

Table 7

White Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular/Aide Class-Size

---

Achievement		Length of Time in Regular/Aide Class-size			
Measures	Grade	1 year	2 years	3 years	4 years
SAT Reading	K	438.88	-----	-----	-----
	1	533.38	504.20	-----	-----
	2	595.38	580.77	582.56	-----
	3	621.21	615.50	612.94	628.09
BSF Reading	K	-----	-----	-----	-----
	1	27.03	23.68	-----	-----
	2	40.86	37.69	38.81	-----
	3	33.17	32.28	32.00	34.14
SAT Math	K	489.53	-----	-----	-----
	1	537.95	517.48	-----	-----
	2	590.32	575.62	584.91	-----
	3	624.10	618.22	614.63	631.91
BSF Math	K	-----	-----	-----	-----
	1	39.91	36.95	-----	-----
	2	53.71	50.57	51.80	-----
	3	51.96	51.45	49.84	53.31

---

regardless of time spent in regular-with-aide class-size, are consistently higher than BSF measures on student subsamples exposed 4 years.

### Research Question (3)

What patterns of differences in white-minority student achievement (reading and mathematics NRT mean scaled scores and CRT average percentage passing) are associated with exposure to small (S), regular (R), and regular-with-aide (RA) class-sizes for one year (grades K, 1, 2, or 3), two years (grades K+1, 1+2, or 2+3), three years (grades K+1+2 or 1+2+3) and four years (grades K+1+2+3)? Having examined the patterns of achievement for minority students and white students separately, Research Question (3) addresses the comparative achievement, or the achievement gap, of white and minority students as a function of class-size type, duration, and incipience of treatment. As in the previous analyses, tabulations are made of each class-size for white and minority subsamples separately (Tables 8-10). However, to determine the possible effect of (S) in preventing the achievement gap, the researcher also cross-tabulated minority (S) and white (R) achievement differences (Table 11). To compare graphically the K, K+1, K+1+2, and K+1+2+3 achievement differences for white (R) and minority (R) and minority (S) subsamples, the researcher charted histograms, one for each achievement measure (Figures 1-4).

Table 8 shows the white-minority achievement gap by grade and length of time in (S). With one exception (BSF Reading, grade 3, 4 years), the 2 years gaps are smallest and the 3 years gaps largest. In grade 3, the second smallest gaps vacillate between 1 year and 4 years length of time spent in (S). The sizes of the gaps are large, ranging from 27.44 points between the grade 1, 1 year and 2 years subsamples on the SAT Reading measure to 6.57 percentage points between the grade 2, 2 years and grade 3, 3 years subsamples on the BSF Math measure.

Table 8

**The White-Minority Achievement Gap: Differences in Stanford Achievement Test (SAT)  
and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time  
in Small Class-Size**

---

Achievement		Length of Time in Small Class-size			
Measures	Grade	1 year	2 years	3 years	4 years
SAT Reading	K	9.91	-----	-----	-----
	1	29.31	17.96	-----	-----
	2	28.44	1.00	16.77	-----
	3	19.36	12.91	24.52	20.65
BSF Reading	K	-----	-----	-----	-----
	1	0.59	0.28	-----	-----
	2	2.85	1.74	3.13	-----
	3	2.87	2.87	5.30	2.52
SAT Math	K	16.64	-----	-----	-----
	1	25.18	20.32	-----	-----
	2	26.55	9.94	19.00	-----
	3	20.46	16.93	25.45	21.49
BSF Math	K	-----	-----	-----	-----
	1	1.49	0.48	-----	-----
	2	2.25	0.45	2.66	-----
	3	5.18	4.93	7.02	5.16

---

Table 9 shows the white-minority achievement gap by grade and length of time in (R). Except for the apparent moderate negative linear relationship between the achievement gap and length of time spent in (R) suggested by the grade 2 and grade 3 outcomes, no clear patterns seem to emerge. Even in the case of the latter, three out of the four achievement measures register a difference score gain after 2 years treatment. In short, the gap between both SAT and BSF measures seems to vary almost randomly.

Table 10 shows the white-minority achievement gap obtained from (RA) subsamples. Clearly, the grade 3, 3 years length of time in (RA) subsamples mean differences were smallest. The average difference between the highest and lowest gap score for the SAT measures was more than 22.00 points, while the average percentage passing difference on the BSF was more than 2.00 points. For each measure, the grade 3, 4 years subsamples gap was either largest or next-to-largest.

Table 11 shows minority (S) and white (R) achievement differences. On every achievement measure, the gaps are smallest at grade K (SAT Reading and Math) or grade 1, 2 years (BSF Reading and Math). Except for SAT Reading, grade 2, 2 years and SAT Math, grade 2, 1 year, the grade 3, 3 years gaps are largest. Uniformly, the grade 3, 4 years achievement gaps are smaller than the grade 3, 3 years gaps and, except for SAT Reading, smaller than the 2 years gaps. The sizes of these gaps are consistent with the other analyses, ranging up to .16 standard deviation and over 2.00 percentage points. Additional results may be obtained by comparing Table 11 with the other gap tables (Tables 8-10) showing the white-minority differences for any one class-size treatment. On every achievement measure and at every grade level, the obtained differences are smaller, sometimes by as much as 23.00 points.

Figures 1-4 graphically illustrate the comparative achievement of white (R) and minority (R) and minority (S) for each achievement measure across one year (K), two years (K+1), three years (K+1+2), and four years (K+1+2+3). Figure 1

Table 9

The White-Minority Achievement Gap: Differences in Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Regular Class-Size

---

Achievement		Length of Time in Regular Class-size			
Measures	Grade	1 year	2 years	3 years	4 years
SAT Reading	K	11.10	-----	-----	-----
	1	35.95	33.20	-----	-----
	2	33.15	32.95	22.33	-----
	3	25.78	28.80	18.38	19.62
BSF Reading	K	-----	-----	-----	-----
	1	2.43	3.75	-----	-----
	2	4.80	6.43	2.82	-----
	3	3.34	4.05	3.05	2.19
SAT Math	K	10.57	-----	-----	-----
	1	27.14	29.85	-----	-----
	2	29.43	35.82	14.74	-----
	3	23.64	24.00	22.92	20.55
BSF Math	K	-----	-----	-----	-----
	1	3.33	4.23	-----	-----
	2	3.56	3.86	2.19	-----
	3	6.08	7.19	5.39	5.05

---

Table 10

The White-Minority Achievement Gap: Differences in Stanford Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by Grade by Length of Time in Reg/Aide Class-Size

---

Achievement		Length of Time in Regular/Aide Class-size			
Measures	Grade	1 year	2 years	3 years	4 years
SAT Reading	K	10.27	-----	-----	-----
	1	40.20	28.87	-----	-----
	2	35.16	38.20	25.76	-----
	3	22.95	20.93	12.48	22.91
BSF Reading	K	-----	-----	-----	-----
	1	2.20	1.75	-----	-----
	2	5.05	6.95	2.53	-----
	3	3.17	2.44	2.03	2.98
SAT Math	K	19.89	-----	-----	-----
	1	28.34	22.90	-----	-----
	2	32.81	43.77	27.76	-----
	3	25.12	22.81	9.83	26.40
BSF Math	K	-----	-----	-----	-----
	1	3.10	3.82	-----	-----
	2	3.80	6.15	2.30	-----
	3	5.75	5.70	2.56	5.85

---



Table 11

**Minority Small-Class and White Regular-Class Subsample Differences in Stanford  
Achievement Test (SAT) and Basic Skills First (BSF) Mean Scaled Scores by  
Grade by Length of Time in Class**

---

Achievement						
Measures	Grade	Length of Time in Class				
		1 year	2 years	3 years	4 years	
SAT Reading	K	5.35	-----	-----	-----	
	1	17.13	6.08	-----	-----	
	2	20.46	10.50	16.91	-----	
	3	13.80	12.79	16.06	14.73	
BSF Reading	K	-----	-----	-----	-----	
	1	0.44	0.67	-----	-----	
	2	2.00	0.94	3.49	-----	
	3	2.18	2.58	4.06	1.95	
SAT Math	K	8.46	-----	-----	-----	
	1	14.04	11.43	-----	-----	
	2	19.72	11.57	16.49	-----	
	3	15.47	19.34	20.83	16.80	
BSF Math	K	-----	-----	-----	-----	
	1	0.82	0.29	-----	-----	
	2	1.33	1.33	2.26	-----	
	3	4.16	4.64	5.33	4.44	

---

shows the results for SAT Reading. At each grade level combination, the gap between white (R) and minority (S) is narrower than that created by the two (Rs). The gap between white (R) and minority (S) achievement is narrowest at grades K and K+1 and grows increasingly larger in K+1+2 and K+1+2+3. In the last two periods, minority (S) and minority (R) achievement is nearly identical. Overall year-to-year achievement gains, however, are substantially higher for all subsamples.

Figure 2 shows the results for BSF Reading. (No BSF tests were administered in grade K.) In grade 1, the minority (S) subsample outscored both white (R) and minority (R) subsamples. In subsequent grades, minority (S) achievement was virtually indistinguishable from minority (R) achievement, while white (R) achievement was comparatively and uniformly higher. For all subsamples, initial surges in achievement from 1 to 1+2 were followed by declines in 1+2+3.

Figure 3 shows the results for SAT Math. This histogram is almost identical to that for SAT Reading (Figure 1). The greatest comparative differences occur in K+1. The difference in minority (R) and (S) is negligible for both K+1+2 and K+1+2+3.

Figure 4 shows the results for BSF Math. As in the case of the histograms for the SAT measures, the BSF histograms are remarkably similar. Initial differences (grade 1) are maximal. In subsequent periods, the white (R) subsamples achieve passing rates several percentage points higher than minority (R) or (S) subsamples. Differences between both minority subsamples are but a few hundredths of a percentage point.

Figure 1

Comparison of SAT Reading Mean Scaled Scores for White (R) and Minority (R) and (S)  
Subsamples That Began in K and Matriculated All Four Years of STAR

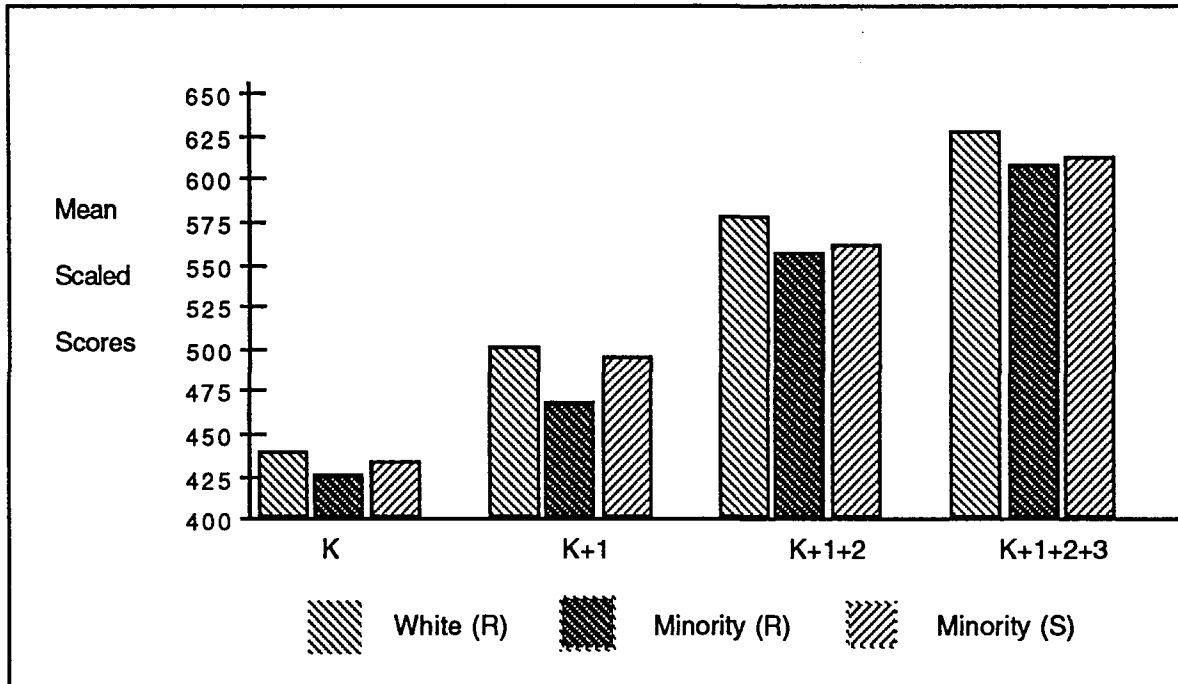


Figure 2

Comparison of BSF Reading Mean Percentage Passing for White (R) and Minority (R) and (S) Subsamples That Began in K and Matriculated All Four Years of STAR

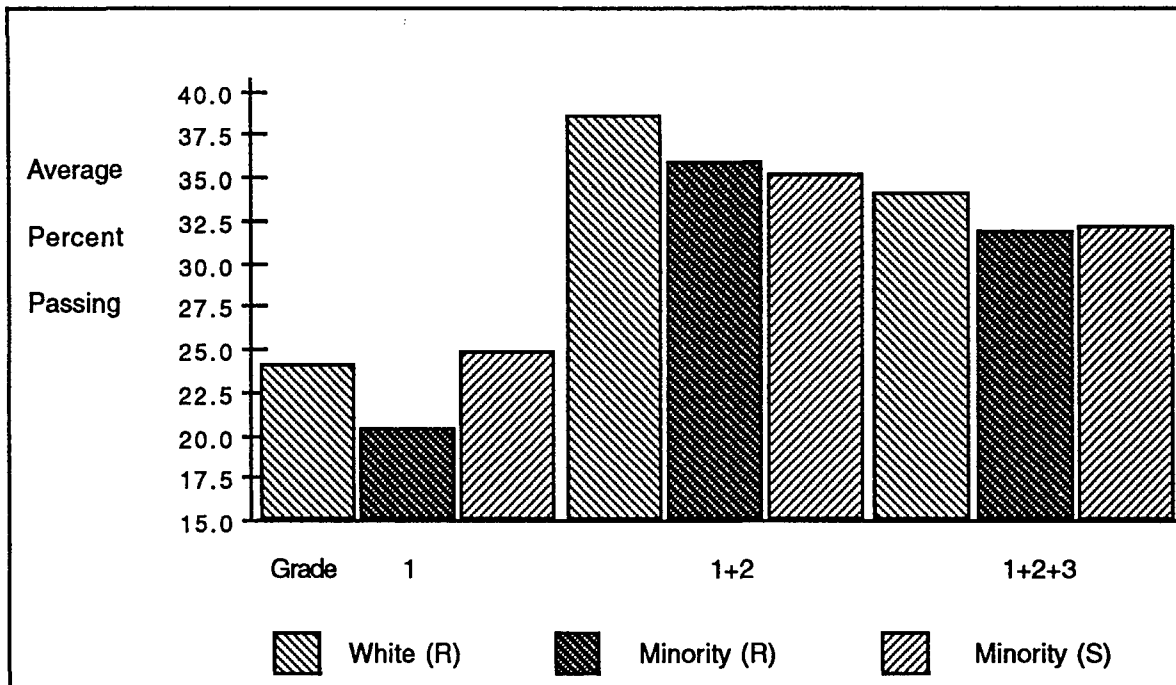


Figure 3

Comparison of SAT Math Mean Scaled Scores for White (R) and Minority (R) and (S) Subsamples That Began in K and Matriculated All Four Years of STAR

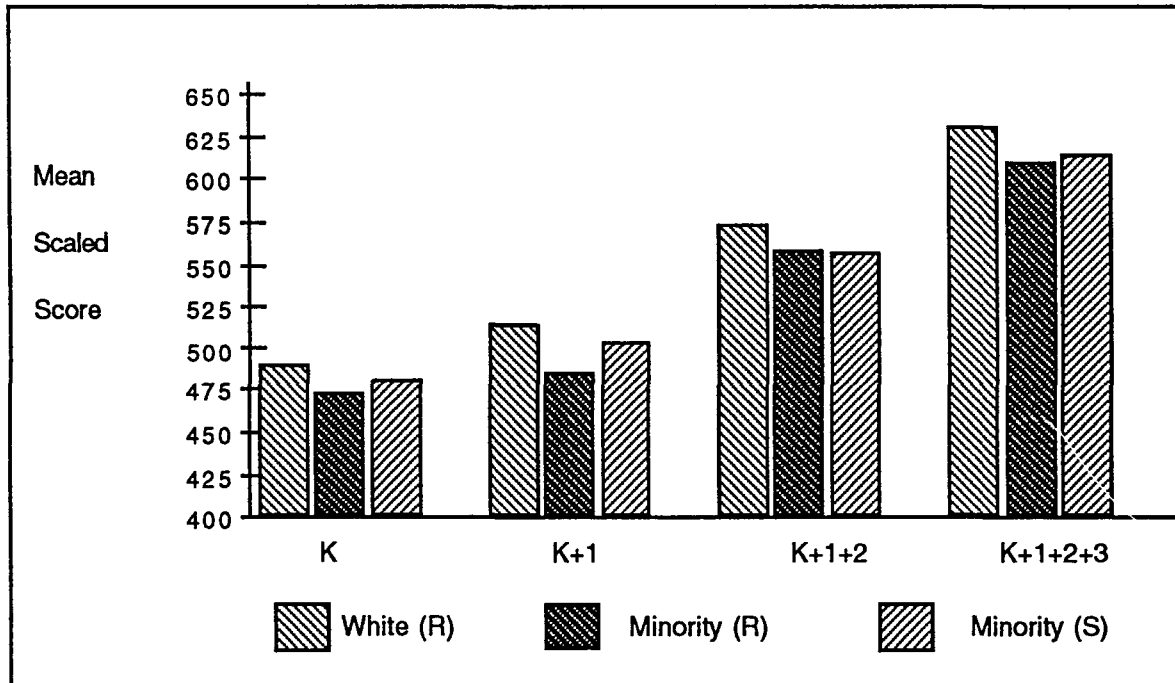
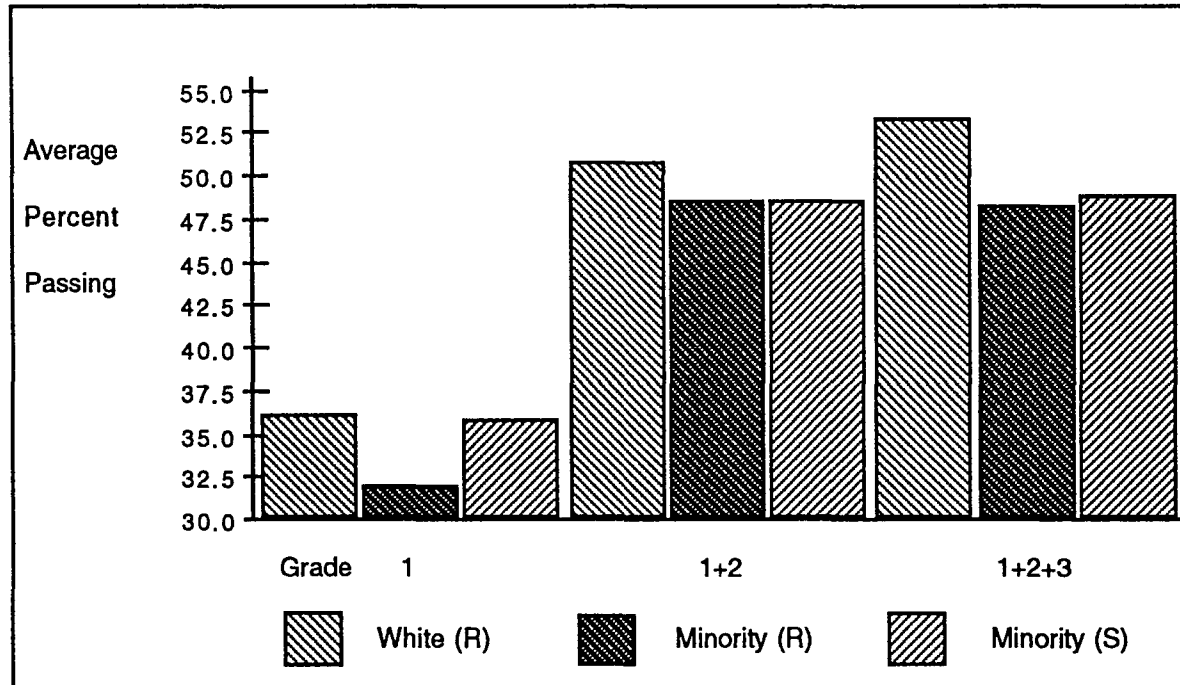


Figure 4

Comparison of BSF Math Mean Percentage Passing for White (R) and Minority (R) and (S) That Began in K and Matriculated All Four Years of STAR



## CHAPTER V

### DISCUSSION

#### Summary of Results

The achievement of minority and white students randomly assigned to Project STAR small (S), regular (R), and regular-with-aide (RA) class-sizes for one, two, three, or four years in grades K, 1, 2, and 3 was analyzed through the tabulation and histography of mean scaled score norm-referenced test (NRT) and criterion-referenced test (CRT) data. The findings are summarized for (a) minority achievement for all years and class-sizes, (b) white achievement for all years and class-sizes, (c) differences between minority and white achievement for all years and class-sizes, (d) achievement differences between minority (S) classes and white (R) classes for all years, (e) achievement differences between white (R) and minority (S) and (R) for K, K+1, K+1+2, and K+1+2+3 subsamples.

#### Minority Achievement for All Years and Class-Sizes

Overall, minority subsamples in (S) tended to obtain higher mean scaled scores on the reading and mathematics areas of the Stanford Achievement Test (SAT)--the NRT component--and to pass the Basic Skills First (BSF) tests--the CRT component-- with greater frequency than either (R) or (RA) subsamples. (As in previous sections of this dissertation, BSF "mean percentage passing" is considered a "score.") Similarly, (RA) subsamples routinely outscored (R) subsamples on all achievement measures except those from the grade 2, 2 years subsamples and the grade 3, 4 years subsamples. Except for the grade 2 BSF measures (regardless of time in treatment), grade 3, 4 years mean scores were typically highest for all class-size treatments. For (S) and (R), a consistent pattern of comparatively high 1 year mean scores with

diminishing 2 years and 3 years scores, followed by the greatest achievement from 4 years subsamples emerged. The pattern for (RA) subsamples, however, differed in that half of the 3 years achievement measures were second only to those of grade 3, 4 years measures. For students who began the (RA) treatment in first grade (grade 3, 3 years subsample), three out of four achievement outcomes were second only to those of students who had begun (RA) in kindergarten (grade 3, 4 years subsample). In terms of effect sizes, maximal differences within (S) and (R) subsamples approached one-half a standard deviation (.49 sigma) and tended to be between 3 years and 4 years time in treatment groups. Obvious in their magnitude, both reading and math BSF outcomes in the (RA) grade 2, 3 years subsamples exceeded those of grade 3, 4 years subsamples by two-thirds of a standard deviation (.67 sigma).

#### White Achievement for All Years and Class-Sizes

The pattern displayed by white subsamples in the (S), (R), and (RA) class-size treatments tended to reiterate that of minority subsamples: Comparatively high 1 year mean scores with diminishing 2 years and 3 years scores were followed by the greatest achievement in 4 years subsamples. Similarly, (S) subsamples tended to outscore (RA) subsamples while (RA) subsamples typically outscored (R) subsamples. The major difference between the minority and white pattern, however, was in the magnitude of the between-treatment subsample means. Differences for white subsamples tended to be compressed relative to differences between minority subsamples, that is, class-size appeared to make a bigger difference for minority students than for white students. Moreover, the tendency of (RA) 3 years achievement outcomes to be surpassed only by (RA) 4 years outcomes failed to be reiterated in the white subsamples. However, the grade 2, 3 year BSF reading score (but not the math score) continued to exceed that of the grade 3, 4 years subsample. In terms of effect sizes, maximal differences across achievement measures averaged



about one-third standard deviation (.34 sigma) for (S) subsamples, while (R) and (RA) subsamples showed greater differences (.43 sigma and .44 sigma, respectively).

#### Differences Between Minority and White Achievement for All Years and Class-Sizes

At any given grade, (S) subsample difference (gap) scores tended to be smallest for the 2 years and largest for the 3 years time in treatment. For (R) subsamples, the gap between both SAT and BSF measures seemed to vary with no clearly discernable pattern. The pattern demonstrated by the (RA) subsamples, however, was consistently marked by gap scores that were lowest for grade 3, 3 years time in treatment. Alternately, grade 3, 4 years subsample gaps were either largest or next-to-largest.

#### Achievement Differences Between Minority (S) Classes and White (R) Classes for All Years

On every achievement measure, the gaps were smallest at grade K, 1 year and grade 1, 2 years. Moreover, grade 3, 4 years achievement gaps were uniformly smaller than grade 3, 3 years gaps and, except for SAT Reading, smaller than the grade 3, 2 years gaps. Grade 2, 2 years gaps, however, were routinely and substantially smaller. Compared to white-minority differences displayed within any one class-size, the gaps that result from white (R) and minority (S) were smaller on every achievement measure and at every grade level.

#### Achievement Differences Between White (R) and Minority (S) and (R) for K, K+1, K+1+2, and K+1+2+3 Subsamples

Results were obtained on the cohort which matriculated synchronistically with the four years of STAR class-size treatment. Histograms for white (R) and minority (R) and (S) on both SAT and BSF achievement measures were remarkably similar. On every measure, the initially large gap between minority (R) and (S) and the small gap between white (R) and minority (S) diminished after one year (BSF measures) or two years (SAT measures). By grade 2, however, achievement differences between

minority (R) and (S) subsamples were negligible while white (R) achievement was routinely and substantially greater than either minority class-size.

#### Conclusions

The purpose of this investigation was to examine patterns of achievement among white and minority students in small, regular, and regular-with-aide class-size treatments over varying lengths of time and at varying start-up years, grades K, 1, 2, and 3. The investigator reasoned that "best" achievement gap reduction intervention would involve delivering treatment in the optimally effective quantity (length of time) and at the optimally effective time (start-up year). The investigator reasoned further that, to be a politically and educationally acceptable goal, small class-size (or any class-size) as a gap reduction intervention must be shown not only to benefit minorities, but not to harm whites. Consequently, the investigator examined absolute (within a group) and comparative (between the two groups) achievement for minorities and whites. To illustrate the effect of no intervention (considered in this context as R) vs. the effect of minority (S) as a gap reduction intervention strategy, the investigator extended the comparative examination to include different class-sizes within and between groups.

#### The Length of Time Issue

Relative to the length of time issue, the results showed that absolute achievement on NRT and CRT measures of both minorities and whites was optimal (highest) for students who had experienced all four years of small-class treatment. Moreover, minority between class-size differences were relatively larger than white between class-size differences, a fact which suggests that the effect of small classes on achievement was of greater benefit to minorities than to whites. This finding supports earlier STAR, LBS, and CHALLENGE research (Finn & Achilles, 1990; Achilles et al., 1992).

The comparative achievement analyses indicated, however, that the differential benefit for minorities did not seem to last. The white-minority achievement gap, smallest after two years time in small-class treatment, appeared to "fade" in subsequent years. (The "fade" issue will be addressed in the next section of this paper.) The same finding was obtained in the minority small-class vs. white regular-class contrast; however, the resulting gaps were uniformly smaller on every measure and at every grade. The diminishing differential effect of small class-size for minorities was graphically apparent in the contrast of that group with white and minority regular-class students that participated in all four years of STAR.

#### The Start-Up Year Issue

When one considers absolute achievement, the question of optimizing the benefit of small class-size by starting treatment in one or another school year highlights an intriguing finding of the present study: Except for the cohort that started in grade K and spent four years time in treatment, one year exposed students score higher than multiple year exposed students. The answer to this anomaly may, in part, be a function of the retention-promotion practices of STAR teachers and the scores of promoted vs. retained students. The percent of grade retentions in STAR was smallest in the (S) condition (7.8 vs. 12.6 in the R and 10.8 in the RA, grade 1) Moreover, students who were promoted tended to score lower (e.g., 422 in S vs. 427 in R on SAT Reading, K). (Percentages and scores are from Word et al., p. 171.) The combination of fewer retainees and lower scores resulted in more "marginal" students matriculating on time with the cohort. Thus when tests were administered at any given grade, more academically capable and older students entering Project STAR for the first time outscored the less capable nonretained students in the original cohort. Over time, this phenomenon results in an apparent "fade" of early intervention benefits (Barnett, 1992). One year treatment means were higher than two and three year

treatment means, however, not just for small-classes but for all class-sizes. A possible explanation is that increased attention was afforded STAR students (and teachers) regardless of class-size treatment. Accompanied by the possibility of higher expectations of and for STAR participants, a Hawthorne effect could have been generated. This observation merits additional investigation.

The comparative achievement pattern analyses examining white vs. minority differences for each class-size found that gap scores tended to be least for the two years in small-class group. Interestingly, this finding held true whether the start-up year was K, 1, or even 2; however, the smallest gaps were registered by the groups that began treatment in grade 1. When minority small-class achievement was compared to white regular-class achievement, the gaps were shown to be least at grade K and grade 1, two years treatment.

#### What is the Optimal Time in Small Class-Size and When Should It Begin?

Consistent with prior STAR, LBS, and CHALLENGE research, the findings of this re-analysis support the following conclusions:

- ( 1 ) Small class-size may be an effective white-minority achievement gap reduction strategy in the primary school years.
- ( 2 ) Benefits accrue initially in greater measure to minorities than to whites.
- ( 3 ) Optimally beneficial (defined as resulting in least white-minority achievement differences) treatment begins no later than grade 1 (preferably in K).
- ( 4 ) Optimally beneficial treatment lasts at least two years, grades K and 1 or 1 and 2.
- ( 5 ) The differential effect of small class-size for minorities appears to "fade" after tow years treatment or in grade 3.

- ( 6 ) As it applies uniquely to small class-size, the "fade" phenomenon in early treatments can be neither confirmed nor contradicted by this re-analysis. To the extent that STAR small-class students were less likely to be retained and that retention in grade is an indicator that a student will drop out of school (Hahn, 1987), the nonretention of minority and white students in Project STAR broadens to life-long the optimization of small-class treatment.

#### Implications of the Results

##### Theoretical Implications

The implications of this study may be conceptualized as both theoretical and practical.

In terms of addressing the theoretical aspect, one might well ask: What has the present research contributed to our knowledge of "gap reduction" strategies for achievement differences between white and minority students in the primary grades of school? Viewed from Cultural Deprivation Theory attributing school failure to deficient home and community conditions of minority youngsters (Lansa & Potter, 1984; Sizemore, 1990), the achievement gains that accrue initially in greater measure to minority than to white small-class students imply that the pernicious effects of minority (read 'economically deprived') culture may be offset in a small-class condition. In part, the nature of language usage among low socioeconomic (SES) minorities may provide an explanation for why this seems to occur. Research has suggested, for example, that children from low socioeconomic homes enter school with restricted language codes using words that are concrete and descriptive as opposed to abstract and analytic (Bernstein, 1962a; 1962b; 1963). Building on Bernstein's work, Achilles (1974) identified low (SES) school children's language as a "communication block" that may act as a verbal barrier between middle class teachers

and lower class students. In a comprehensive study of elementary school reading curricula, Marzano (1991) analyzed and identified basic words in grades K-6, the acquisition of which would "level the field" for low SES (disproportionately minority) students who entered school with a smaller vocabulary relative to middle and high SES students. Consequently, if small classes in this study provided minority and culturally disadvantaged children with enhanced opportunity for more and cognitively richer verbal interaction with the teacher (and one would logically assume they did), then the results tend to support Cultural Deprivation Theory generally and language deprivation theory specifically.

Beyond the small-class condition itself, the findings of this study relative to the incipience and duration of "treatment" yield important theoretical implications. This re-analysis disclosed that optimally beneficial (defined as resulting in least white-minority achievement differences) treatment begins no later than grade 1 (preferably in K) and lasts at least two years. Why then and why that long? Although the Effective Schools research and literature reviewed in Chapter II (e.g., Denton & Davis, 1988; Jennings, 1992; Murphy, 1988; Ramey, 1993) does not answer the question directly, most strategies successful in reducing the white-minority achievement gap substantiate the need for early intervention. Bryson (1993) has marshalled paleobiological evidence to suggest that the human brain is prewired to experience a dramatic surge in learning from about age four and that, most importantly, the brain's maps are being established then for all future learning. The implication is that, if schools are to effect positive differences in children's capacity for cognitive or academic achievement, then the intended intervention must occur as close to kindergarten as possible. In a practical illustration of this concept, Crain, Mahard, and Narot (1982) have shown that achievement gains are seen for black students attending desegregated schools from kindergarten or first grade, but not for

black students whose desegregated education began in later school years. Similarly, Jennings (1992) has reported that the efficacy of Reading Recovery and Success for All in enhancing disadvantaged minority student achievement is founded upon implementation no later than first grade. Moreover, if Bryson's (1993) research is correct, the young child's brain, "glow[ing] like a hickory fire on a cold winter night (p. 16)," may have acquired in two years most of what is biologically possible to acquire toward obtaining parity with more cognitively advantaged peers. The findings of this study seem to suggest that for small class-size intervention, the same conclusions obtain--treatment should occur no later than age five or six and will likely yield optimal results after two years.

#### Practical Implications

What are the practical implications of the results? These may best be described as recommendations for policymakers and practitioners. First, although a cost-benefit analysis exceeds the scope of this dissertation, the expense of implementing small classes in the primary years of school must be weighed against the cost of remediation (e.g., Chapter 1 reading programs) in later years, the effectiveness of which is questionable. As a report of the Perry Preschool Project recently disclosed (Lewis, 1993), the benefits of early childhood intervention programs tend to yield lifelong dividends (e.g., higher levels of education, higher economic status, and greater social responsibility) for self and society. For preschool programs, the latest research suggests that \$7.16 is the eventual saving to the public for every dollar invested. Although exact figures are impossible to estimate at this time, small class-size, too, may be a cost efficient public-school option. Absent contradictory evidence, small classes should be implemented in grades K and 1 whenever possible.

Second, education policymakers and practitioners must act in ways that reflect sensitivity to the concerns of ethnic and racial minorities. To the extent that espousing Cultural Deprivation Theory may be viewed as the oppression of a minority culture by a majority one, it is ethically wrong. Gardner (1991, p. 53), for example, reports that the Kaluli people of New Guinea see babies as helpless creatures who neither understand nor are capable of speech. Rather than speaking to the children, Kaluli mothers speak "for them." Yet the family unit is strong with the generations cohabiting until death parts them. Despite the viability and grace of Kaluli practices, the tendency to label such a culture "deprived" vis-a-vis our predominantly white, Anglo-Saxon, protestant, middle-class American culture is great. Clearly, desegregation and compensatory education are among the more beneficial, socially constructive consequences of advancing Cultural Deprivation Theory. However, educators and education policymakers must not permit application of the model in everyday life to minimize or obliterate the good and moral qualities inherent in all cultures and peoples.

#### Recommendations for Future Study

This study was intended to be an investigation of small-class as a "gap reduction" strategy for achievement differences of white and minority students participating in Project STAR. Accordingly, this researcher described within and between group achievement patterns associated with time in treatment and start-up year of treatment for grades K, 1, 2, and 3. The findings generated by the copious quantity of data analyzed has given birth to as many questions as answers. This investigator recommends for future study the following:

- (1) Using the comparative analyses of the present study, investigate in greater detail (RA) as gap reduction strategy. The (RA) achievement patterns often differed in significant ways from either the (S) or (R) patterns for both whites and



minorities. One might extend the comparative analyses, for example, to examine achievement differences between minority (RA) classes and white (R) classes for all years.

(2) Investigate possible causes for the differential effects of class-size, time in treatment, and start-up year on the SAT achievement measures vs. the BSF measures. These measures sometimes varied in directly opposite ways. To the extent that the BSF tests were purported to assess TN public school instructional outcomes, one must wonder why, at a rate of 50 percent, successive years' mean percentage of whites in (R) and (RA) passing the BSF declined as the more generic SAT outcomes rose.

(3) Employ least-squares analysis to determine the point at which achievement outcomes, shown to be greater for (S) at every grade level and on every measure, for (R) and (RA) would "catch up" with (S). Near equivalent SAT (but not BSF) gains were associated with all groups' matriculation from one grade to the next. To quantify the average differences and project these into the future would provide educators and policymakers with information useful in selecting from among competing gap reduction alternatives.

(4) Investigate possible causes for the consistency of high white and minority one year achievement outcomes for all grades, achievement measures, and class-sizes. This finding does not reconcile with the "fade" effect wherein long-term gains appear to be washed out simply because of the comparison group makeup. If academically "marginal" (S) students were promoted at greater rates than (R) or (RA) students, then why were the one year scores of (R) and (RA) students also comparatively high? (The question of high one year achievement scores is currently under investigation by this writer. Using a data set with fewer missing cases, the SAT results for (a)

minority (S), (b) white (S), (c) white-minority gap (S), and (d) minority (S) and white (R) are shown in Appendix C.)

(5) Replicate Project STAR with even greater numbers of students and class-sizes. Despite its enormity, the database from which this re-analysis drew was too small to permit an examination of the data by location or socio-economic status. As David Armor of George Mason University has demonstrated, whether a child is poor and lives in an urban setting (demonstrably related phenomena) is of greater effect on achievement than the color of his skin (Judson, 1993). Perhaps time in treatment and start-up year effects are different for small class-sizes in urban as contrasted with suburban locations. Such analyses await a larger database.

(6) Continue to track STAR small class-size students as they matriculate through school. Eventually, it may be possible to conduct a cost-benefit analysis to determine if the initial investment yielded dividends for the individual participants as well as for society. Such information would be quite valuable for education policymakers.

**BIBLIOGRAPHY**

- Achilles, C. M. (1974, Fall). "Oh, I thought you meant." New Jersey School Development Council Research Bulletin, 20(1), 30-32.
- Achilles, C. M. (1987). Problem solving is not problem finding! Paper presented at the NOVA University Summer Institute, San Diego.
- Achilles, C. M., & DuVall, L. (1982). Interim evaluation report: Project SHAL. St. Louis Public Schools: St. Louis, MI.
- Achilles, C. M., & DuVall, L. (1983). Some preliminary notes of an effective schools report: Project SHAL. (ERIC Document Reproduction Service No. ED 247 653).
- Achilles, C. M., Nye, B. A., Zaharias, J. B., Fulton, B. D., & Wallenhorst, M. P. (1992, August). Class size is a reasonable policy alternative for educational excellence--or is it? Paper presented at the National Conference of Professors of Educational Administration, Indiana State University, Terre Haute, IN.
- Achilles, C. M., Nye, B. A., Zaharias, J. B., Fulton, D., & Bingham, S. (1993, August). Prevention or remediation: Is small class a reasonable treatment for either? Paper presented at the National Conference of Professors of Educational Administration, Palm Springs, CA.
- Ackerman, P. L. (1989). Review of the Stanford Early School Achievement Test, Third Edition. In J. C. Conoley and J. J. Kramer (Eds.), The tenth mental measurements yearbook (pp. 866-867), Lincoln, NE: The Buros Institute of Mental Measurements.

- Allington, R. L., & Johnston, P. (1989). Coordination, collaboration, and consistency: The redesign of compensatory and special education intervention. In R. E. Slavin, N. L. Karweit, N. A. Madden (Eds.), Effective programs for students at risk (pp. 320-354), Boston: Allyn and Bacon.
- Armor, D. J. (1972, Summer). The evidence on busing. Public Interest, 28, 90-126.
- Bain, J. G., & Herman, J. L. (Eds.). (1990). Making school work for underachieving minority students: Next steps for research, policy, and practice. New York: Greenwood Press.
- Barnett, S. W. (1992, Spring). Benefits of compensatory preschool education. Journal of Human Resources, 279-312.
- Bernstein, B. (1962a). Linguistic codes, hesitation phenomena, and intelligence. Language and Speech, 5(1).
- Bernstein, B. (1962b). Social class, linguistic codes, and grammatical elements. Language and Speech, 5(4).
- Bernstein, B. (1963, June). Social structure, language, and learning. Educational Research, 3.
- Bobbett, G., French, R., & Achilles, C. (1991, August). Role of school district report cards: Why do we bet on .25 with .75 running loose? Paper presented at the annual meeting of the National Council of Professors of Educational Administration, Fargo, ND.
- Bodmer, W. F. & Cavalli-Sforza, L. L. (1970, October). Intelligence and race. Scientific American, 19-29.
- Brookover, W. B. et al. (1979). School social systems and student achievement: Schools can make a difference. New York: Praeger.

- Bryson, J. (1993). Leadership, the new science, and a new learning theory for schools for the information age. Unpublished manuscript, University of North Carolina at Greensboro, School of Education, Greensboro, NC.
- Carter, T. P., & Chatfield, M. C. (1986). Effective bilingual schools: Implications for policy and practice. American Journal of Education, 95, 200-234.
- Carpenter, D. (1989). Review of the Stanford Achievement Test, Eighth Edition. In J. C. Conoley and J. J. Kramer (Eds.), The tenth mental measurements yearbook (pp. 861-863), Lincoln, NE: The Buros Institute of Mental Measurements.
- Chase, C. (1980, November). Three views in the complex controversy on the origin of intelligence. Phi Delta Kappan, 6(11), 217-218.
- Christie, J., & Johnsen, E. (1983). The role of play in social-intellectual development. Review of Educational Research, 53(1), 93-115.
- Coleman, J. S. (1972, Summer). Coleman on 'The Coleman report.' Public Interest, 28, 127-128.
- Coleman, J. S., Campbell, E. O., Holson, C. J., McPartland, J., Mood, A. M., Weinfield, F. D., & York, R. L. (1966). Equality of Educational Opportunity. Washington, D. C.: U. S. Office of Education, National Center for Educational Statistics.
- Conant, J. B. (1961). Slums and suburbs. New York: McGraw-Hill.
- Crain, R. L., & Mahard, R. E. (1978). Desegregation and black achievement. Law and Contemporary Problems, 42(3), 17-56.
- Crain, R. L., Mahard, R. E., & Narot, R. (1982). Making desegregation work. Cambridge, MA: Ballinger.
- CTB/McGraw-Hill (1989). Technical bulletin #2. New York: CTB Macmillan/McGraw-Hill.

- Cummins, J. (1986). Empowering minority students: A framework for intervention. Harvard Educational Review, 56, 18-35.
- Denton, S. E., & Davis, L. E. (1988). Collaborative intervention approaches for at-risk youth: The Pittsburg experience. (ERIC Document Reproduction Service No. ED 319 816).
- Diegmüller, K. (1993, March 24). Conservative-leaning think tanks putting imprint on education policy, Education Week, 12(26), 1,36.
- D'Onofrio, W. D. (1983). Test scores in New Castle County, DE.: Before and after busing. (ERIC Document Reproduction Service No. ED 236 260).
- Edmonds, R. (1979a, September). Effective Schools for the urban poor. Educational Leadership, 37(1), 15-24.
- Edmonds, R. (1979b). Some schools work and more can. Social Policy, 9, 28-32.
- Edmonds, R. (1981). Making public schools effective. Social Policy, 12, 56-60.
- Edmonds, R. (1982, December). Program of school improvement: An overview. Educational Leadership, 40, 4-11.
- Eysenck, H. J., & Kamin, L. (1981). The intelligence controversy. New York: John Wiley.
- Finn, J. D., Fulton, D., Zaharias, J., & Nye, B. A. (1989/1992). Carry-over effects of small classes. Peabody Journal of Education, 67(1), 75-84.
- Finn, J. D., & Achilles, C. M. (1990). Answers and questions about class size: A statewide experiment. American Educational Research Journal, 27(3), 557-577.
- Folger, J. (1989/1992). Project STAR and class size policy. Peabody Journal of Education, 67(1), 1-16.
- Gallagher, M. P., & Lanese, J. (1991). Reading study: 1989-90. (ERIC Document Reproduction Service No. ED 339 014).

- Garcia, E. (1988). Effective schooling for language minority students (New Focus No. 1). Washington, DC: National Clearinghouse for Bilingual Education.
- Garcia, E. (1991). Education of linguistically and culturally diverse students: Effective instructional practices. (ERIC Document Reproduction Service No. ED 338 099).
- Gardner, H. (1991). The Unschooled mind: How children think & how schools should teach. New York: Basic Books.
- Garner, R. (1990). When children and adults do not use learning strategies: toward a theory of settings. Review of Educational Research, 60(4), 517-529.
- Gerard, H. B. (1983). School desegregation: The social science role. American Psychologist, 38, 869-877.
- Gerard, H. B., & Miller, N. (1975). School desegregation. New York: Plenum.
- Glass, G. V., Cahen, L. S., Smith, M. L., & Filby, N. N. (1982). School class size: Research and policy. Beverly Hills: Sage.
- Glass, G. V., & Hopkins, K. D. (1984). Statistical methods in education and psychology (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Glassman, P., & Roelle, R. J. (1990). Closing the achievement gap: How comprehensive changes in program and curriculum design can improve achievement for all and narrow the gap between minority student achievement and Caucasian student achievement. (ERIC Document Reproduction Service No. ED 328 588).
- Gold, N. C. (1985). Competency testing for limited-English-proficient students. (ERIC Document Reproduction Service No. ED 273 151).
- Goldsboro City Schools. (1987). Responsive Early Childhood Education Program (RECEP). (ERIC Document Reproduction Service No. ED 309 847).

- Good, T. L., Grouws, D. A., Mason, D. A., Slavings, R. L., & Cramer, K. (1990). An observational study of small-group mathematics instruction in elementary schools. American Educational Research Journal, 27(4), 755-782.
- Gould, S. J. (1981). The mismeasure of man. New York: W. W. Norton.
- Hahn, A. (1987, December). Reaching out to America's dropouts: What to do? Phi Delta Kappan, 69(4), 256-263.
- Hamburg, D. A. (1992). Today's children: Creating a future for a generation in crisis. New York: Times Books.
- Harrington, M. (1963). The other America. Baltimore: Penguin Books.
- Hathaway, W. (1990, October). Monitoring and improving school learning. In J. G. Bain & J. L. Herman(Eds.), Making schools work for underachieving minority students: Next steps for research, policy, and practice. New York: Greenwood Press.
- Haycock, K. (1990). Equity, relevance, and will. In J. G. Bain & J. L. Herman (Eds.), Making schools work for underachieving minority students: Next steps for research, policy, and practice. New York: Greenwood Press.
- Haycock, K., & Brown, P. R. (1984). Excellence for whom? A report from the planning committee for the Achievement Council. (ERIC Document Reproduction Service No. ED 277 787).
- Haycock, K., & Navarro, M. S. (1988). Unfinished business: Fulfilling our children's promise. (ERIC Document Reproduction Service No. ED 299 025).
- Hodgkinson, H. (1991, September). Reform versus reality. Phi Delta Kappan, 73(1), 9-16.
- Hodgkinson, H. L. (1992, October). A demographic look at tomorrow. Paper presented at the annual fall conference of Personnel Administrators of North Carolina, Raleigh, NC.



- Howard, J. & Hammond, R. (1985, September). Rumors of inferiority. The New Republic, 17-21.
- Ipka, V. W. (1992). The Norfolk decision: The effects of converting from a unitary educational system to a dual educational system upon academic achievement. (ERIC Document Reproduction Service No. ED 346 583).
- Isonio, S. (1990). LAUSD student performance on the 1988-89 Scholastic Aptitude Test: Description and comparative analysis. (ERIC Document Reproduction Service No. ED 331 918).
- Jaeger, R. M. (1982). An iterative structured judgment process for establishing standards on competency tests: Theory and application. Educational Evaluation and Policy Analysis, 4(4), 461-476.
- Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B., & Michelson, S. (1972). Inequality: A reassessment of the effect of family and schooling in America. New York: Harper and Row.
- Jennings, J. M. (1992, November). Closing the achievement gap: A model for success. Paper presented at the Southern Regional Council of Educational Administrators Annual Conference, Georgia State University, Atlanta, GA.
- Jenson, A. R. (1969). How much can we boost I.Q. and scholastic achievement? Harvard Educational Review, Winter, 1-123.
- Jenson, A. R. (1981). Straight talk about mental tests. New York: Free Press.
- Jones, L. V. (1984). White-black achievement differences: The narrowing gap. American Psychologist, 39(11), 1207-13.
- Jones, L. V. (1987). Achievement trends for black school children, 1970-1984. (ERIC Document Reproduction Service No. 285 944).

- Judson, G. (1993, February 20). Poverty, not race, is linked to Hartford School Failures. The New York Times, p. 31.
- Keeves, J. P. (1988). Longitudinal research methods. In J. P. Keeves (Ed.), Educational research, methodology, and measurement: An international yearbook. Oxford: Pergamon Press.
- Kozol, J. (1991). Savage inequalities: Children in America's schools, Boston: Crown.
- Kratochwill, T. R. (1978). Single subject research: Strategies for evaluating change. New York: Academic Press.
- Lanese, J. (1988). Longitudinal reading parity study 1979-1987. (ERIC Document Reproduction Service No. ED 306 545).
- Lansa, P. J., & Potter, J. H. (1984, January). Building a bridge to span the minority-majority achievement gap. Paper presented at the National Conference on Desegregation in Postsecondary Education, Durham, NC.
- Lewis, A. C. (1993, June). The payoff from a quality preschool. Phi Delta Kappan, 74(10), 748-749.
- Levine, D. U., & Lezotte, L. W. (1990). Unusually effective schools: A review and analysis of research and practice. University of Wisconsin, Madison (WI): The National Center for Effective Schools Research & Development.
- Lezotte, L. W. (1989, August). Base school improvement on what we know about effective schools. American School Board Journal, 18-20.
- Lemann, N. (1986, June). The origins of the underclass. The Atlantic Monthly, 31-55.
- Levin, H. (1990). The educationally disadvantaged are still among us. In J. G. Bain & J. L. Herman (Eds.), Making schools work for underachieving minority students: Next steps for research, policy, and practice. New York: Greenwood Press.

- Lucas, T., Hentz, R., & Donato, R. (1990). Promoting the success of Latino language minority students: An exploratory study of six high schools. Harvard Educational Review, 60, 315-334.
- Madden, N. H., Slavin, R. E., Karweit, N. L., Dolan, L., & Wasik, B. A. (1991). Success for all: Multi-year effects of a schoolwide elementary restructuring program. Johns-Hopkins University, Baltimore, MD: Center for Research on Effective Schooling for Disadvantaged Students.
- Marks, J. L. (1985). The enrollment of black students in higher education: Can declines be prevented? (ERIC Document Reproduction Service No. ED 264 817).
- Marzano, R. J. (1991). The analysis and identification of basic words in grades K-6. (ERIC Document Reproduction Service No. ED 337 758).
- Martinez, M. E. (1985). Educational opportunity for underrepresented minority students in California: Comparisons with other states. (ERIC Document Reproduction Service No. ED 262 127).
- Mayo, S.T. (1989). Review of the Stanford Achievement Test, Eighth Edition. In J. C. Conoley and J. J. Kramer (Eds.), The tenth mental measurements yearbook (pp. 861), Lincoln, NE: The Buros Institute of Mental Measurements.
- McGhee, J. D. (1985). The black family today and tomorrow. The State of Black America 1985, New York: National Urban League.
- Mitchell, D. E., Beach, S. A., & Badarak, G. (1989/1992). Modeling the relationship between achievement and class size: A re-analysis of the Tennessee Project STAR data. Peabody Journal of Education, 67(1), 18-74.
- Moe, J. L., Nacoste, R. W., & Insko, C. A. (1981). Belief versus race as determinants of discrimination: A study of Southern adolescents in 1966 and 1979. Journal of Personality and Social Psychology, 41, 1031-1050.

- Moll, L. (1988). Educating Latino students. Arts, 64, 315-324.
- Mullis, I., Owen, E. H., & Phillips, G. W. (1990). Accelerating academic achievement - America's challenge: A summary of findings from 20 years of NAEP. (ERIC Document Reproduction Service No. ED 325 500).
- Nye, B. A., Achilles, C. M., Zaharias, J. B., Fulton, B. D., & Wallenhorst, M. P. (1992a, April). Five years of small-class research: Student benefits derived from reduced student/teacher ratios. Paper presented at the American Educational Research Association, San Francisco, CA.
- Nye, B. A., Achilles, C. M., Zaharias, J. B., Fulton, B. D., & Wallenhorst, M. P. (1992b). Small is far better. Paper submitted for MidSouth Educational Research Association Outstanding Paper Competition, Knoxville, TN.
- Ogbu, J. U. (1992). Understanding cultural diversity and learning. Educational Researcher, 21(8), 5-14.
- Ornstein, A. C. (1982). The education of the disadvantaged: A 20-year review. Educational Research, 24(3), 197-211.
- Orum, L. S. (1990). Making education work for Hispanic-Americans: Some promising community-based practices. In J. G. Bain & J. L. Herman (Eds.), Making schools work for underachieving minority students: Next steps for research, policy, and practice. New York: Greenwood Press.
- Parrenas, F. Y., & Parrenas, C. S. (1990). Cooperative learning, multicultural functioning and student achievement. (ERIC Document Reproduction Service No. ED 337 540).
- Parsons, D. (1993, January 31). Equity 2000 aims at closing mathematics achievement gap. AASA Leadership News, p. 6.

- Pate-Bain, H., Achilles, C. M., Boyd-Zaharias, J., & McKenna, B. (1992). Class size does make a difference. Phi Delta Kappan, 74(3), 253-256.
- Pease-Alvarez, C., Garcia, E., & Espinosa, P. (in press). Effective schooling in preschool settings: A case study of LEP students. Early Childhood Research Quarterly.
- Pelavin, S. H. (1988). Changing the odds: Factors increasing access to college. New York: College Entrance Examination Board.
- Peterson, N. L. (1987). Early intervention for handicapped and at-risk children: An introduction to early childhood special education. Denver, CO: Love Publishing Company.
- Peterson, K., & McCue, L. (1990). The North Carolina school improvement project: Resource notebook. University of Wisconsin, Madison (WI): The National Center for Effective Schools Research & Development.
- Piaget, J. (1962). Play, dreams, and childhood. New York: Norton.
- Pinnell, G. S., DeFord, D. E., & Lyons, C. A. (1988). Reading Recovery: Early intervention for at-risk first graders. (ERIC Document Reproduction Services No. ED 303 790).
- Ramey, M. (1984). Seattle classroom observation system manual. Seattle: Seattle Public Schools Research and Evaluation Office.
- Ramey, M. (1992, April). Classroom characteristics related to ethnic achievement gap reduction. Paper presented at the annual meeting of the American Education Research Association, San Francisco.
- Ramey, M., Hillman, L., & Matthews, T. (1982, March). School characteristics associated with instructional effectiveness. Paper presented at the annual meeting of the American Educational Research Association, New York.

- Ravitch, D. (1983). The troubled crusade: American education 1945-1980. New York: Basic.
- Riessman, F. (1962). The culturally deprived child. New York: McGraw-Hill.
- Rothman, R. (1993, January 20). Achievement in math improves, NAEP data find. Education Week, pp. 1, 23.
- Sandia National Laboratories. (1991). Perspectives on education in America: Annotated briefing-third draft. Systems Analysis Department, Albuquerque, NM.
- Scherer, M. (1993). On Savage Inequalities: A conversation with Jonathan Kozol. Educational Leadership, 50(4), 4-9.
- Shapson, S. M., Wright, E. N., Eason, G., & Fitzgerald, J. (1980). An experimental study of the effects of class size. American Educational Research Journal, 17, 141-152.
- Schuhman, A. M. (1987). Improving the quality of teachers for minority students. (ERIC Document Reproduction Service No. ED 294 954).
- Sizemore, B. (1985). Pitfalls and promises of effective schools research. The Journal of Negro Education, 54(3).
- Sizemore, B. (1989). Curriculum, race and effective schools. In H. Holtz, I. Marcus, J. Dougherty, J. Michaels, & R. Peduzzi (Eds.), Education and the American dream. Granby, MA: Bergin and Gavey Publishers.
- Sizemore, B. (1990). Effective education for underachieving African-Americans. In J. G. Bain & J. L. Herman (Eds.), Making schools work for underachieving minority students: Next steps for research, policy, and practice. New York: Greenwood Press.
- Slavin, R. E. (1983). Cooperative Learning. New York: Longman.

- Slavin, R. E. (1990). Research on cooperative learning: Consensus and controversy. Educational Leadership, 47(4), 52-54.
- Slavin, R. E., Karweit, N. L., & Wasik, B. A. (1993). Preventing early school failure: What works. Educational Leadership, 50(4), 10-18.
- Staples, R. (1987). Social structure and Black family life: An analysis of current trends. Journal of Black Studies, 17(3), 267-87.
- Southern Regional Education Board (SREB). (1986). Measuring student achievement. Atlanta, Author.
- St. John, N. H. (1975). School desegregation: Outcomes for children. New York: John Wiley.
- Sternberg, R. J. (1988). The triarchic mind: A new theory of human intelligence. New York: Viking.
- Sylva, K., Bruner, J. S., & Genova, P. (1974). The role of play in the problem-solving of children 3-5 years old. In J. S. Bruner (Ed.), Play--its role in development and evolution. New York: Basic Books.
- The Psychological Corporation (1985). Stanford Achievement Test series. Technical data report. Orlando, FL: Harcourt, Brace, Jovanovich.
- TN Department of Education (1987). STAR criterion referenced test. Manual for test administrator. Grade 1. Nashville, TN: State Education Department.
- Willis, S. (1993, February). Closing the achievement gap. ASCD Update, 35(2), 3,8.
- Word, E., Johnston, J., Bain, H., Fulton, B., Zaharias, J., Lintz, N., Achilles, C. M., Folger, J., & Breda, C. (1990). Student/teacher achievement ratio (STAR): Tennessee's K-3 class size study. Final report and final report summary. Nashville, TN: Tennessee State Department of Education.
- Wright, W. J. (1992). The endangered Black male child. Educational Leadership, 49(4), 14-16.

Zafirau, S. J. (1986). An analysis of 1982 to 1986 reading comprehension achievement results for adversely affected students in the Cleveland public schools. (ERIC Document Reproduction Services No. ED 277 806).



APPENDIX A: Achievement Measures Total n of Cases by Time in Treatment by Class-Size by Race

			K	1	2	3	K+1	1+2	2+3	K+1+2	1+2+3	K+1+2+3
	S	W	1187	1231	1193	1311	175	16	130	83	154	831
		M	552	592	598	602	84	29	95	62	91	309
SAT	R	W	1359	1500	1302	1176	182	40	195	79	206	553
Read		M	646	934	746	639	145	55	142	60	116	197
	RA	W	1357	1492	1427	1452	169	60	182	57	278	762
		M	687	644	790	808	91	51	213	59	147	245
	S	W		1257	1240	1269	173	19	125	84	150	804
		M		595	636	586	83	31	95	65	89	299
BSF	R	W		1551	1345	1173	187	41	197	74	196	547
Read		M		932	787	645	145	58	143	64	115	199
	RA	W		1550	1467	1408	174	62	181	59	267	739
		M		650	816	815	90	53	214	60	146	248
	S	W	1204	1271	1189	1338	181	40	131	83	158	847
		M	558	596	597	599	84	28	97	61	91	306
SAT	R	W	1370	1550	1467	1408	187	40	194	78	211	559
Math		M	661	946	748	637	149	56	142	61	116	196
	RA	W	1374	1570	1424	1493	180	60	185	56	289	779
		M	703	654	788	803	91	52	213	59	147	242
	S	W		1255	1266	1296	172	20	126	87	153	819
		M		586	635	587	80	30	95	65	89	300
BSF	R	W		1543	1368	1202	183	44	201	77	202	558
Math		M		932	787	643	144	58	143	64	114	199
	RA	W		1552	1512	1449	175	63	186	61	279	756
		M		644	817	814	90	53	213	60	146	248

APPENDIX B: Achievement Measures Standard Deviations by Time in Treatment by Class-Size by Race

			K	1	2	3	K+1	1+2	2+3	K+1+2	1+2+3	K+1+2+3
	S	W	33.52	57.92	46.74	38.84	65.71	57.28	38.60	50.34	38.89	37.36
		M	29.07	48.06	38.34	36.97	49.89	50.96	35.45	41.66	34.62	36.23
SAT	R	W	31.60	54.13	45.80	36.90	55.33	50.33	36.27	48.15	37.30	37.25
Read		M	27.38	44.32	35.24	33.10	39.06	37.38	29.62	37.43	36.33	32.67
	RA	W	31.77	55.22	45.11	37.77	61.49	48.37	36.27	48.65	38.65	35.64
		M	29.83	41.75	38.32	34.13	44.31	39.59	32.54	40.30	36.44	32.79
	S	W		4.64	6.96	5.36	6.42	11.20	5.53	9.32	5.01	4.92
		M		5.04	7.37	6.87	6.42	8.01	6.72	9.23	7.60	6.60
BSF	R	W		5.06	7.30	5.58	6.23	10.26	5.67	8.96	6.07	5.14
Read		M		6.10	8.48	6.76	6.56	8.81	6.64	9.91	7.39	6.17
	RA	W		4.93	6.95	5.73	6.38	8.93	5.79	8.41	6.33	4.95
		M		5.58	8.76	7.04	6.19	10.22	6.66	9.30	6.73	6.67
	S	W	48.19	44.44	47.06	39.73	53.67	58.93	36.32	54.95	38.25	39.82
		M	50.44	38.34	37.18	37.10	39.57	36.86	30.72	41.12	32.20	38.64
SAT	R	W	45.55	40.39	42.98	37.97	43.20	50.50	35.84	44.07	42.83	36.58
Math		M	50.10	38.11	37.47	36.74	32.65	36.38	35.36	39.96	36.92	38.92
	RA	W	44.37	41.83	41.82	39.36	48.72	43.12	35.20	50.08	38.54	39.19
		M	45.68	38.54	39.93	35.08	38.04	40.54	36.08	43.38	36.23	34.12
	S	W		5.18	6.60	7.54	8.90	7.07	7.66	9.75	7.46	6.93
		M		6.12	6.76	9.75	8.64	5.90	9.77	8.94	10.66	9.42
BSF	R	W		5.52	6.90	8.07	8.51	10.37	7.87	7.27	9.85	6.80
Math		M		7.63	7.60	9.46	9.85	7.50	9.58	9.87	10.20	8.58
	RA	W		5.27	6.40	8.18	7.33	11.16	7.89	8.05	9.22	7.04
		M		6.86	8.34	10.05	8.45	11.32	9.68	8.89	10.13	9.65

**APPENDIX C: Tables of Additional Analyses of Star Data Pertinent to Small Class as a "Gap Reduction" Strategy for Achievement Differences**

---

**Minority Stanford Achievement Test (SAT) Mean Scaled Scores by Grade by  
Length of Time in Small Class-Size Obtained from "Unfiltered" Data Set**

---

**Achievement**

Measures	Grade	Length of Time in Small Class-size			
		1 year	2 years	3 years	4 years
SAT Reading	K	433.79	-----	-----	-----
	1	498.24	513.74	-----	-----
	2	565.83	565.66	575.28	-----
	3	602.79	606.19	594.55	612.78
SAT Math	K	479.56	-----	-----	-----
	1	511.88	524.42	-----	-----
	2	563.28	562.58	572.39	-----
	3	610.89	604.39	594.90	613.63

---

**White Stanford Achievement Test (SAT) Mean Scaled Scores by Grade by Length  
of Time in Small Class-Size Obtained from "Unfiltered" Data Set**

---

**Achievement****Measures****Length of Time in Small Class-size**

	<b>Grade</b>	<b>1 year</b>	<b>2 years</b>	<b>3 years</b>	<b>4 years</b>
<b>SAT Reading</b>	<b>K</b>	<b>443.69</b>	-----	-----	-----
	<b>1</b>	<b>523.60</b>	<b>542.73</b>	-----	-----
	<b>2</b>	<b>587.66</b>	<b>584.93</b>	<b>604.67</b>	-----
	<b>3</b>	<b>611.12</b>	<b>619.10</b>	<b>619.07</b>	<b>633.43</b>
<b>SAT Math</b>	<b>K</b>	<b>496.20</b>	-----	-----	-----
	<b>1</b>	<b>538.08</b>	<b>548.42</b>	-----	-----
	<b>2</b>	<b>582.68</b>	<b>582.42</b>	<b>599.83</b>	-----
	<b>3</b>	<b>615.19</b>	<b>621.32</b>	<b>620.35</b>	<b>635.12</b>

---

**The White-Minority Achievement Gap: Differences in Stanford Achievement Test (SAT)**

**Mean Scaled Scores by Grade by Length of Time in Small Class-Size**

Obtained from "Unfiltered" Data Set

---

<b>Achievement</b>		<b>Length of Time in Small Class-size</b>			
<b>Measures</b>	<b>Grade</b>	<b>1 year</b>	<b>2 years</b>	<b>3 years</b>	<b>4 years</b>
<b>SAT Reading</b>	<b>K</b>	9.90	-----	-----	-----
	<b>1</b>	25.36	28.99	-----	-----
	<b>2</b>	21.83	19.27	29.39	-----
	<b>3</b>	8.33	12.91	24.52	20.65
<b>SAT Math</b>	<b>K</b>	16.64	-----	-----	-----
	<b>1</b>	26.20	24.00	-----	-----
	<b>2</b>	19.40	19.84	27.44	-----
	<b>3</b>	4.30	16.93	25.45	21.49

---

**Minority Small-Class and White Regular-Class Subsample Differences in Stanford  
Achievement Test (SAT) Mean Scaled Scores by Grade by Length of Time in Class**  
Obtained from "Unfiltered" Data Set

---

**Achievement**

Measures	Grade	Length of Time in Class			
		1 year	2 years	3 years	4 years
SAT Reading	K	5.28	-----	-----	-----
	1	18.36	18.96	-----	-----
	2	21.17	14.04	42.22	-----
	3	16.50	12.79	16.06	14.73
SAT Math	K	8.78	-----	-----	-----
	1	16.92	15.58	-----	-----
	2	25.25	16.66	20.39	-----
	3	7.28	19.34	20.83	16.80

---