Order Number 9419161

The effect of class size on retention in the primary grades: Implications for educators and policy makers

Harvey, Barbara Ann Hawks, Ed.D.

The University of North Carolina at Greensboro, 1993

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

THE EFFECT OF CLASS SIZE ON RETENTION

IN THE PRIMARY GRADES: IMPLICATIONS

FOR EDUCATORS AND POLICY MAKERS

by

Barbara Hawks Harvey

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

C. M. Achilles

Dissertation Advisor

HARVEY, BARBARA HAWKS, Ed.D. The Effect of Class Size on Retention in the Primary Grades: Implications for Educators and Policy Makers. (1993) Directed by Dr. Charles Achilles. 123 pp.

This post hoc study analyzed data collected through the STAR Project in Tennessee. This project followed students in early primary (k-3) from 1985 through 1989 in order to study the effects of class size on achievement. Lasting benefits studies are on-going. The current study focused on class size as it affected the achievement scores of students retained in kindergarten and first grade and those not retained in those grades. Demographics for both groups were compared and a portrait of the retained kindergartner and retained first grader was drawn.

A literature review of retention and class size was conducted and presented. The sample of retained students in kindergarten and grade one was taken from the STAR database. Demographic and achievement data were obtained for the sample and analyzed using an analysis of variance.

Findings from this study showed that once a child was retained in kindergarten or grade one, small class size offered no remedial or preventive measures. It was also found that over the period of tracking, retained students scored consistently lower than their non-retained counterparts.

The retained kindergartner was typically a white male in rural schools. The retained first grader possessed these same characteristics. Both students came from low socio-economic backgrounds. A noted difference existed between these two grades within the percentages of males and females retained.

This study offers support for the supposition that retention offers no benefits in the realm of increasing achievement. Further, it suggests that small classes are beneficial to boosting test scores. Additional research is needed in the area of class size and retention.

Approval Page

This dissertation has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Dissertation Advisor C. M. [illegible]

Committee Members Joseph E. Bryson

Dale L. [illegible]

[illegible]

11/17/93

Date of Acceptance by Committee

11/1/93

Date of Final Oral Examination

ACKNOWLEDGEMENTS PAGE

It is with much appreciation that I acknowledge the support and aid extended me by my advisor Dr. Charles Achilles and the members of my committee--Dr. Dale Brubaker, Dr. Joseph Bryson, and Dr. Rita O'Sullivan.

With heart-felt gratitude, I recognize the encouragement and support my mother, sister, brother, son, and husband, my best friend, and my principal and staff have so freely given over the past three years. Thank you.

I would also like to recognize the Center of Excellence for Research in Basic Skills at Tennessee State University, Nashville, Tennessee for database use and technical assistance. Thanks also to personnel in the Knox County, Tennessee schools for confirming grade retention within a portion of the STAR data. Results expressed in this study do not necessarily reflect the position of the Center or its staff.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Grade retention, despite decades of use and abuse, simply does not work. In fact, it never did. Originally, American schools were basically ungraded. Students moved through their education, advancing as they mastered the content, the system operating much as does today's outcome-based education. It was not until the late 19th century and the arrival of industrial-age thinking that the one-room schoolhouse succumbed to German influence and the majority of American children began climbing from one grade to the next via age (Goodlad & Anderson, 1963). Problems began to arise. Many youngsters mastered the curriculum with ease, while others had some degree of difficulty, and still others were not at all successful. Discipline presented a dilemma as did effectiveness of instruction when all students were moved forward to the next grade despite their level of skill. The very structure of schools was beginning to wobble, and it certainly could not be beneficial to promote children to the next level if they lacked the skills necessary to be successful. Not to worry. A simple solution lay on the horizon--retention or non-promotion.

Non-promotion would alleviate such problems as these and so by 1900, failure rates were at 50%, with the majority of those retained being in the primary grades (Doyle, 1989).

One difficulty with this simple solution was that retention failed to remedy the very ills which it was intended to treat. Non-promotion would allow students who had not reached the school's standards an extra year to catch up; they might catch up with the subject, but they would never catch up with their peers. These children would, thus, avoid the emotional trauma associated with low achievement and the schools would, in turn, avoid the discipline and instructional problems associated with this group. Not so. Research as early as 1911 reported that evidence did not support retention as being more beneficial than grade promotion for students with academic and/or adjustment difficulties. In fact, none of the 44 studies conducted from 1911 to 1973 could offer confirmation that retention accomplished its purpose (Jackson, 1975).

While Jackson's (1975) review of these studies was heralded as valuable, it was the meta-analysis review of retention literature carried out by Holmes and Matthews (1984) that drew the most attention. Their results related that retained students achieved .44 SD lower in achievement than did the control group and that they also were significantly lower than the promoted students in social adjustment (.27 SD), emotional adjustment (.37 SD), and behavior (.31 SD), as well as on measures of self-concept and attitude toward school. Concluding from these results, Holmes and Matthew stated:

> Those who continue to retain pupils at grade level do so despite cumulative research evidence showing that the potential for negative effects consistently outweighs positive outcomes. Because this cumulative research evidence consistently points to negative effects of nonpromotion, the burden of proof legitimately falls on proponents of retention plans to show there is compelling logic indicating success of their plans when so many other plans have failed (p.232).

Smith and Shepard (1987) also found that retention was not justifiable, citing it as just one more component in education that has a long history of common wisdom but that does not work. Norton (1983) stated that retention aided neither pupil achievement nor personal adjustment. Byrnes and Yamamoto (1983) decried retention as harmful to self-esteem, and Frymier (1989) has said that the heart attack victim has a better chance of surviving than does the child facing grade retention. Yet, despite the overwhelming body of research that finds grade retention to be ineffective and even harmful in many cases, educators continue its use.

PROBLEM

The pendulum often swings from one extreme to the other in educational reform. Student retention has not escaped this phenomenon. Since the early 1800's, retention in grade has been a common practice. At the turn of the century, the average retention rate for all grades was 16%. By 1930, social scientists began questioning the value of retaining students and suggested that there might be negative effects from retention. The retention rate dropped to approximately 5% in the 1940's with social promotion being anointed as one alternative to retention. In the 1960's social promotion became widespread. Critics were quick to notice declining achievement scores and emphasized a concern with promoting students who lacked the necessary skills to move ahead with their peers.

The pendulum swung once again toward retention in the 1980's; the Gallup Poll (1986) showed that 72% of the US citizenry favored stricter grade-to-grade promotion standards. Consequently, retention rates climbed toward 7% annually. It has been estimated that 5.6 million students in the United States, 14% of the total 40 million school population, have repeated a grade during the past 12 years (Frymier, 1989). The January 1990 Policy Brief from the Center for Policy Research in Education estimated that the overall expenditure for retention in the US is $10 billion per year. By ninth grade, 50% of all US students have failed at least one grade or have dropped out of school (Shepard & Smith, 1989). Statistics relate that even one grade retention increases the risk of high school dropout from 10% to 40% (Safer, 1983) while some studies say that two years of retention will increase the chances of dropping out to 90% (Hahn, 1987).

Throughout the years, it has been more common to retain students in kindergarten and first grade than in later grades (Rose, Medway, Cantrell & Marus, 1983). Delidow's (1989) longitudinal study of 166 students indicates that 75% of all retention occurs before 3rd grade. The reasons for retaining younger children are plentiful: political pressures to maintain high standards, insistence on teacher accountability, development of more homogeneous first grade classes, age, physical size, social maturity, school and behavioral problems, parental emphasis on more academic skills, readiness levels, ad infinitum (Nason, 1991; Doyle, 1989; Shepard and Smith, 1986; Uphoff, 1985; Langer et al, 1984). In some school districts, as many as 60% of kindergartners are judged unready for 1st grade and so are retained or placed in transitional programs (Shepard & Smith, 1988). These developmental or junior kindergartens often become dumping grounds for children who do not fit into a homogeneous kindergarten (Billman, 1988, p.10). A larger number of males and low-income students are also found in these programs (Billman, 1988; Charlesworth, 1989). Such inequities coupled with these staggering rates have today's educators, parents, and legislators once again questioning the worth of grade-level retention.

Research on the value of retention has not been a carefully guarded secret, though many educators and most policy makers appear never to have been acquainted with any of it. Since the early 1900's, over 100 studies have been conducted on the subject of retention (Medway, 1985). Such studies over the past 80 years have concluded that: 1)retention does not increase learning; students who are promoted tend to learn more than students of like ability who were retained, 2)retention does not increase reading readiness for most students, 3)retention does not increase socialization skills, and 4)retention tends to promote discipline problems (Norton, 1983). Otto and Estes (1960) reviewed research from the 1930's and 1940's in this area, concluding that:

> Repetition of grades has no special educational value for children; in fact, the educational gain of the majority of nonpromoted students subsequent to their retention is smaller than that of their matched age mates who were promoted. Similarly, the threat of failure has no appreciable positive effect on the educational gain of the threatened...(pp.4-11).

In 1975, the retention literature was again reviewed, this time by Jackson. His conclusions were similar to those of Otto and Estes: "There is no reliable body of evidence to indicate that grade retention is more beneficial than grade promotion for students with serious academic or adjustment difficulties" (627). Research in the 1980's was no more favorable toward grade retention than it had been a decade earlier. Yamamoto (1980) looked at the emotional effects of retention on children. He discovered that children rated going blind and losing a parent as the only two events more stressful than being retained. Byrnes' (1989) interviews with children related that youngsters associated retention with punishment for being bad or failing to learn. Holding students back has a negative impact on social adjustment and self-esteem, retained students citing such terms as "bad," "sad," and "embarrassed" with repeating a grade (Rose et al., 1983). Nor does retaining students increase the homogeneity of classrooms (Bossing & Brien, 1979; Haddad, 1979).

In addition to the negative emotional effects of retention, studies discredited the contention that retention improved academic achievement. Students who repeat a grade will achieve less than students who are promoted; 40% of retained students learn less than at-risk students promoted to 1st grade and only 20% to 35% acquire new content knowledge (Bossing & Brien, 1980; May & Welch, 1984; Rose et al., 1983; Shepard & Smith, 1989). Holmes and Matthew (1984) reviewed 650 reports, covering 50 years of research, on grade retention and found that retained students scored significantly lower than promoted students on outcome measures. Holmes (1988, 1989) established that 54 of the 63 studies that he used in a meta-analysis related negative effects when students were retained. This led to his deduction that the damaging effects of retention were no longer questionable, but fact. Few practioners have given heed to such conclusive evidence against retention.

Many people continue to assert that high standards will be maintained through retention of those who have not garnered the skills necessary for the next grade level. Cumulative research shows that this supposition is false. Others believe that promotion is a reward for accomplishments. Doyle (1989) reminds us that every child must be granted the chance to be educated in the learning environment that best provides for his needs; educational opportunity is not a privilege but a right. Thus, recognizing that retention does not produce the meritorious effects intended and that students have the right to the best education they can achieve, educators cannot continue to place blame for failure on the child. They must, instead, replace retention with appropriate alternatives.

Literature on alternatives to retention has increased greatly over the past few years. A majority of the research emphasizes benefits of intervention in the regular classroom for at-risk students. Learning problems should be diagnosed and prescriptions drafted and implemented (Norton, 1990, 206). Lieberman (1980) and Shepard and Smith (1990) suggest that multi-disciplinary teams do in-depth analyses of students who are inadequate or severely deficient in basic skill acquisition. These students then advance to the next grade with Individualized Educational Plans. Recycling students through the same programs which were originally inappropriate for them will only guarantee that the programs are equally inappropriate for them the second time around and less interesting (Cunningham & Owens, 1976, 29). Those who speak against retention advocate promotion with such interventions as peer tutoring and crossage peer tutoring, summer programs, mainstreaming, cooperative learning, attention to learning styles, individualized instruction, special instructional programs on weekends and during vacation, remediation before and after school, parent-help programs (Hartley, 1977; Texas Education Agency, 1987; Bredekamp & Shepard, 1989; Sklarz, 1989; Marshall, 1991). Unlike retention, these options have a research base signifying positive effects. In addition to in-class programs, there are separate alternatives to promotion with remediation. Included are nongraded, multi-aged programs much like those of the first American schools, developmentally appropriate curriculum taught by teachers properly prepared to deliver it, curriculum based on more current learning theory from cognitive and constructivist psychology, and use of smaller classes (Wertsch, 1985; Byrnes & Yamamoto, 1986; Connell, 1987; Resnick, 1987; Charlesworth, 1989). The most often selected alternatives to remediation are increased remedial instruction and smaller classes (Byrnes & Yamamoto, 1986).

To date, there have been few documented efforts to implement any of these alternatives to retention. Seemingly, there is a multitude of reasons: little recognition that a problem exists, little interest in solving the problem, oversimplification of solutions, high stakes accountability, and cost factors. Literature on retention addresses each of these. Retention often equals dropouts. Remedial programs and smaller classes are well planned solutions. Educators are beginning to recognize that the current testing practices are insufficient for measuring achievement. Cost of retention is more than three times the cost of high quality remedial services for a year; compare $3000 to $800 (Allington, 1988 in Norton, 1990, 206). Interestingly, policy makers and practioners have only to read the literature to be confronted by both problem and remedy.

Use of smaller classes is both an alternative to retention and a reasonable remediation step for students who have been retained. Much of the literature in this area confirms that smaller classes do make a difference in students' achievement and development (Glass & Smith, 1978; Glass, Cahen, Smith, & Filby, 1982; Achilles, Bain, & Finn, 1990; Bain, Achilles, Zaharias, & McKenna, 1992). While many of the studies on class size effect have been criticized for lacking such traits as randomness, longitudinal nature, and large scale size, there is at least one study which remediates these weaknesses. Tennessee's state legislature funded a $12 million, four-year study beginning in 1985 called STAR (Student/Teacher Achievement Ratio) which analyzed student achievement and development in three types of classes: small (13-17 students per teacher), regular (22-25 students per teacher), and regular classes with a teacher's assistant. Finn and Achilles (1990) noted: "This research (STAR) leaves no doubt that small classes have an advantage over larger classes in reading and mathematics in the early primary grades" (573). Teachers who participated in this project found that they were better able to identify students' needs, provide more individualized instruction, and cover more material more effectively in the smaller classes. With such positive effects from smaller classes, is it probable that one alternative to retention is smaller classes?

In summary, the problem with retention lies in the fact that, despite a multitude of studies proving that retention is not beneficial, educators and policy makers continue to employ it as a common practice. Students continue to be retained yearly under the guise of higher standards. Those same students continue to fall further behind and many eventually become dropouts. Neither our society nor our economy can continue to lose so much money, so much man power as is lost due to the deleterious practice of retention.

PURPOSE

Although retention has been around for over a century, researchers have discovered, since its inception, that it is a practice with little merit. Numerous alternatives to non-promotion have been offered by innovative educators and researchers. Still, an increase in retention is occurring today as a result of the trend in stricter policies regarding promotion. How can policy and practice operate contrary to the significant body of research which has been available for 80 years?

This study will serve multiple ends. First, it will seek to address class size as an option to retention. Concurrently, the database will be used to determine whether smaller class size will have a positive effect on the achievement scores of those students previously retained. The study will look at at pupils retained and then placed into three types of classes--S, R, and RA. Comparisons will be made between STAR and the state average of retention rates.

The second purpose of this study will be to suggest other options to retention. The literature is replete with research-based alternatives, many of which were unknown at the time retention was considered to be a panacea to the schools' ills. Among the many are cooperative learning, learning styles, cognitive learning theories, developmentally appropriate curriculum, and peer tutoring. These components and others will be detailed in the final chapter of this study.

The third objective will be to add to the relevant body of research on alternatives to retention. This research will draw the attention of education policy makers and practitioners. In so doing, alternatives cited herein will help displace retention as a common practice.

QUESTIONS

Two questions will be addressed in this study through use of the STAR database. What does the retained kindergartner and retained first grader look like in relation to his non-retained peers? If a retained student is subsequently placed in either a small class, regular class, or regular class with an assistant, what are the relative differences in achievement for retained students? Does class size serve as an alternative to or remediation for retention? Additionally, alternatives to retention and policy implications will be attended to based on conclusions from this study.

GENERAL METHODOLOGY

This will be a post hoc study, relying primarily on new analyses of data collected from Project STAR in Tennessee. Project STAR followed students in early primary (K-3) from 1985 through 1989 in order to study the effects of class size. (Lasting benefits studies are on-going.) All Tennessee districts were asked to participate; 42 of the 140 districts were selected with 79 elementary schools in those districts providing sites for STAR. The project included 17 inner-city, 16 suburban, 8 urban, and 39 rural schools. In 1985-1986, there were 6325 kindergartners, with 127 small classes, 103 regular classes, and 98 regular classes with an aide. In 1986-1987, 7103 first graders made up STAR's population. (Tennessee did not require kindergarten in 1985-86 which partially accounts for the larger number of pupils in first grade.) In grade one there were 124 small classes; 115 regular classes, and 108 regular-with-an-aide classes. Students and teachers were randomly assigned to each of the three class types (small, regular, regular with an aide).

New pupils entering the STAR sites were randomly assigned to one of the three class-size conditions.

The main focus of Project STAR was on student achievement. Measurement was based on appropriate forms of the Stanford Achievement Test (SAT) for grades K-3, Tennessee's Basic Skills First (BSF) Criterion Tests (grades 1-3), a test tied closely to the state's curriculum objectives. The primary unit of data collection was the student, but the class (class average) was the unit of analysis. The primary analysis consisted of multivariate tests of mean differences between and among the groups being analyzed. The study concluded that students in small classes made higher scores on the Stanford Achievement Test and on the Basic Skills First Test in all four years in all locales. Results were both educationally and statistically significant.

The present study of retention-in-grade issues related to class size will deal with students in Project STAR who were retained in kindergarten and students retained in first grade. Results based on the achievement tests originally used in STAR will be used to determine whether small class size offers a remedial effect on retained students. An ANOVA will be used in the analysis.

LIMITATIONS AND DELIMITATIONS

The limitations of this study come largely from using the STAR database. Most of this database is already established and so can only offer information previously collected. Additional constraints may be imposed because this study will be conducted as part of a larger, on-going research effort that may limit resources and impose time requirements.

The study is limited to students who were retained in kindergarten and first grade in Tennessee in 1985-86 and who subsequently (1986-87) entered the STAR kindergarten and grade one samples and were randomly assigned to one of the three STAR conditions (S, R, RA). By accepting information provided by school personnel who checked a sample of students who were identified as retained, another limitation is imposed. Additionally, once a student is retained twice, he is lost from the STAR database.

Another limitation stems from the possibility of minor error due to the process for selecting the sample. While the 253 students retained in kindergarten were known, the 1152 students identified as being retained in first grade were estimated based on a pilot and the age of retained kindergartners. Having determined that this dissertation will be based on a post hoc study dictates the sample to be used. While it may have been desirable to include students not in Project STAR, this was not feasible and so becomes a delimitation.

DEFINITION OF TERMS AND ABBREVIATIONS USED IN THIS STUDY

EARLY PRIMARY: grades kindergarten through third (K-3)

NONPROMOTION: see retention

REGULAR-SIZE (REGULAR) CLASS (R): 22-26 students per teacher; $\bar{X}$=24

REGULAR-WITH-AIDE CLASS (RA): 22-26 students per teacher with a full-time aide; $\bar{X}$=24

RETENTION/
RETENTION IN GRADE/
(also nonpromotion): failure to be promoted to the next higher grade; requirement that an elementary student repeat a grade level through a second school year; also referred to as held back, repeating, or failure

SMALL-SIZE (SMALL) CLASS (S): 13-17 students per teacher; $\bar{X}$=15

STAR: Student Teacher Achievement Ratio Project funded by the Tennessee legislature from August, 1985 through August, 1989 to determine class-size effects on pupil achievement and development in early primary grades. Forty-two local school districts were involved with teachers and students being randomly assigned to small, regular, and regular classes with an aide. Lasting Benefits Studies are currently on-going (1989-1993) using the original STAR study as baseline.

STAR DATABASE: a computerized record of STAR's population (n=7100) tracked by individual identification numbers and including demographic data, test data, and such things as attendance, discipline, teacher and principal data, etc.

SIGNIFICANCE OF STUDY

With the trend in competency-based education comes a renewed interest in the effects of retaining students. This study may help determine whether class size is a worthwhile alternative to retention and/or an effective remediation step for once-retained pupils by using the STAR database to compare achievement scores among students in S, R, and RA classes. This study should overcome many of the weaknesses of earlier studies on retention through use of the data collected during Project STAR (a randomly-assigned student-teacher sample) and by analysis of the longitudinal results (grades 1-3).

There have been a multitude of studies on both class size and retention, but many are fraught with problems. Project STAR and this study have been designed to avoid many of the weaknesses of the previous research. Jackson (1975) found that the 44 retention studies that he analyzed were of three different design types. He concluded that no results could be drawn from two of these designs because of biases. Type I studies compared retained pupils who were having difficulties with promoted students who were not, and type II studies failed to control for possible improvement resulting from causes other than retention itself.

As with retention studies, early class-size studies also suffered weaknesses. Methodological problems and design dissimilarities were common. Lack of randomization, the short duration of some treatments, insufficient sample size, failure to account for other classroom factors, and use of the pupil rather than the class as the unit of analysis were among the problems. STAR avoided these pitfalls. The strengths of Project STAR lie in its true experimental nature, something not often achieved in education. STAR possesses randomization, longitudinal analyses, and an adequate sample. The class (class average) was used as the unit of analyses with control for such factors as teacher-effect and student interactions within the class. The database continues to offer a sound research basis for the Lasting Benefits Study (LBS).

This study capitalized on the strengths of Project STAR to answer previously stated questions about retention and class size. The study can provide an important body of research that illustrates the effects of retention and alternatives to non-promotion. Additionally, policy implications are discussed as are a number of options to retention.

ORGANIZATION OF STUDY

This dissertation is presented in five chapters with the general headings of 1)introduction, 2)literature review, 3)methodology, 4)presentation of data, and 5)summary, conclusions, and implications. Chapter I provides a brief introduction about retention and class-size issues. The problem statement addresses the need for alternatives to retention and is followed by the purpose of this study and the research questions to be addressed. Chapter I also includes a summary of the study's methodology, limitations and delimitations, definitions of terms and abbreviations, and the significance of the study. Chapter II, review of previous research and literature, includes two sections: a literature review followed by a review of the research. The chapter provides a historical review of retention as well as a review of the major literature in the area. A general synopsis of class-size studies includes a focus on Project STAR. The chapter concludes with a review of the need for options to retention, raising the question of whether small classes will be an effective alternative and whether it will remediate the problems identified as requiring retention.

Chapter III describes both the methodology of the STAR study and of this study. Sufficient detail is provided to enable a person to replicate the post hoc study. The chapter presents information on population and sample, design, data analysis plan, and also explains data collection beyond the data normally obtained for STAR and\or LBS.

Chapter IV is a compilation of the data collected during research, the analysis of the data and an explanation of its significance. The final chapter, Chapter V, provides a summary of the findings, conclusions, discussion of

CHAPTER II

REVIEW OF THE LITERATURE

THE HISTORY OF RETENTION

Retention has washed the educational shores in waves of popularity since its origin. A review of the literature on retention shows the topic to have been of keen interest since the early 1900's. Although retention has been widely used for nearly a century, the efficacy of this practice remains questionable and its usefulness controversial. Educators and researchers alike hold highly emotional views on the issue. Advocates of retention cite the need for standards while its critics hold that those same standards are not achieved by retaining students. The research review overwhelmingly finds in favor of the latter group and has done so since the first days of retention studies.

By 1840, elementary education had been divided into eight grade levels (Bossing & Brien, 1980). Henry Barnard began the crusade in 1838 to transform America's one-room schoolhouses into a system of graded classes based on the Prussian model. Horace Mann, John Pierce, and others contributed to implementation of the plan so that within a two-year period, a system of gradation existed throughout the United States (Tyack, 1974).

Problems began to arise concomittantly, with teachers observing that the homogeneous groups that were desired did not appear. Standard courses of study and mandatory examinations evolved. Some children did not possess the same academic skills as their classmates. Some students were not as emotionally or socially ready as their peers to move to the next grade level, despite their age. A solution was sought and retention identified as the proper treatment for addressing the problems of slow learners and immature youngsters.

As the new system evolved, questions arose and criticism of grade retention policies surfaced. W.T. Harris reduced the rigidity of the system by regrouping those students experiencing difficulties at the end of each six weeks, and the presidents of Harvard and the University of Chicago called for more flexible school organization to support unique abilities. John Dewey, along with others, also challenged the established system, developing experimental models to displace gradation and retention. Reviewing the early research regarding elementary school retention, Saunders (1941) summarized:

> From the evidence cited, it may be concluded that nonpromotion of pupils in elementary schools in order to assure mastery of subject matter does not accomplish its objectives. Children do not appear to learn more by repeating a grade but experience less growth in subject matter achievement than they do when promoted (p.29).

Still, graded schools continued to grow as did retention rates. With it would grow the body of research against both.

By the end of the 19th century, 70% of all students in any one year were affected by repetition of grade (Karweit, 1991). When such figures became available through the media, the public became as concerned about nonpromotion as were many educators. "Educational scientists" labeled grade retention as promoting waste and failure. Numerous research studies ensued. Among the first was a study by Leonard Ayres which led him to write in 1909:

> Under our present system there are large numbers of children who are destined to live lives of failure. We know them in the schools as the children who are always a little behind intellectually, and a little behind in the power to do. Such a child is the one who is always "it" in the competitive games of childhood. (cited in Tyack, 1974, p.199).

This thinking was soon joined by a shift in psychology which underlined youngsters' social and emotional well-being. Studies by Sandin (1944) and Goodlad (1954) revealed that non-promoted children tended to lack confidence and in general were more insecure than promoted children. The tide was beginning to turn. During the 1930's, with its value deemed harmful, retention practices fell into disuse. Educators began to adopt social promotion policies which pushed academically-based policies aside. From 1918 until

1952 there was a decline in over-age students at each grade level. Yet, homogeneous classes were maintained through grouping and tracking. Dropout rates continued to soar.

Social promotion enjoyed popularity for nearly 30 years but could not withstand the launch of Sputnik and the decrease in national standardized achievement test scores. The public began to view social promotion as the offender primarily responsible for America's academic decline, and mastery learning and criterion reference testing came into vogue. A Nation at Risk (1983) increased attention to standards and advancement to the next grade. States across the country began to look for ways to improve public relations. The solution was espied as testing. Minimal competency testing was promptly anchored to exit requirements in order to assuage the public and try to insure the acquisition of adequate academic skills upon graduation. The 1990's have seen an increase in the number of states instituting this requirement; today, 40 states require competency testing for graduation. And so once again, retention rides high on the wave of favor.

Instituting more strenuous academic requirements has meant that school districts have implemented stronger retention policies. In turn, cumulative rates of retention are currently as high as they were before social promotion. Many states have annual retention rates of seven percent;

variance in cities and states ranges from two percent to 20%. Atlanta's retention rate was four times as high in 1981 as in 1980 after implementation of competency testing. Many states have also implemented testing after specific grade levels, which will certainly contribute to the failure rate where these tests are used as exiting requirements.

WHY STUDENTS ARE RETAINED

Two themes form the basis for retention policies: student immaturity or adjustment difficulties and low achievement. Advocates of retention view an extra year at the same grade level as a "gift of time," giving students the opportunity to mature and to be exposed again to material they did not understand the first time around. Gesell (1982) and Ames (1966, 1980) advocated testing to determine the child's developmental age, which should then be used to place the child, not his chronological age. Thirty states reported that they use academic readiness tests prior to kindergarten in some districts and 43 reported that some districts use academic readiness testing prior to first grade (Schultz, 1989). Stringer (1960) stated that the best predictor of an individual's future learning rate is his past learning rate, finding in his study that retention seemed to be more helpful than harmful. Learning is seen as linear, occurring in stages, and fear of failure is seen as a strong motivator by these supporters.

Critics of retention turn to cognitive psychology in an attempt to discredit the retentionist's theory of linear learning which stems from the reductionist's theory of learning. Constructivists propose that learning is most meaningful in context and that it is not necessary for children to progress step by step. On the contrary, cognitive psychologists believe that lower skills are incorporated into more advanced levels of learning. They view learning as a constructive process, focusing on the process of thinking rather than on the end product. Cognitive researchers agree that meaningful learning is reflective, constructive, self-regulated (Branford & Vye, 1989; Davis & Maher, 1990; Marzano et al., 1988). Glasser (1991) summarized this change in thinking:

> Given the growing body of information about human competence and performance, the emphasis in theories of learning has shifted from the accumulation of facts and their reinforcement, to the structure and coherence of knowledge and its accessibility in problem solving and reasoning (p.28).

This shift has also caused the focus on assessment to change. Paper-and-pencil tests are recognized as insufficient to measure thought processes just as testing prior to third grade is deemed inappropriate. Scientific knowledge underlying readiness assessment is explicit in that none of the existing measures is sufficiently accurate to justify removing children from their normal peer group (Shepard & Smith, 1986). Numerous reviewers have reported that the Gesell School Readiness Tests, which are recommended for screening children, do not meet the standards of the American Psychological Association for validity, reliability, or normative information (Kauffman,1985; Shepard & Smith, 1985), and the Metropolitan Readiness Tests would fare only slightly better if used for the same purpose.

When such tests are used, a large number of identification errors occur; it is not possible to make highly accurate assessments of school readiness. Early testing policies are part of the high-standards syndrome with 40 states offering developmental kindergartens and transitional programs in at least some of their schools. Retention rates couched in terms of transitional programs and pre-kindergarten programs often stem from such a practice. Studies reveal a broad range of retention rates in this area: district-level rates in California in 1985-1986 vary from zero to 50%; approximately 8% of kindergarten students in Georgia were retained in 1988; districts in Delaware retain between zero and 11% of their kindergartners and between 8% and 28% of their first-graders, and Boston retained 6.4% of its kindergartners and 19.4% of its first-graders in 1987 (Schultz, 1989). There are obviously a number of educators who put a lot of stock in the old adage, "If at first you don't succeed, try, try again."

Such folk wisdom also encapsulates the apparent belief that the threat of repeating a grade serves as a strong motivator. No study to date has been able to offer support for this supposition. In fact, one study (Otto & Melby, 1935) revealed that students who were told at the year's beginning that they would be promoted regardless of the work that they did exhibited no difference in behavior or quality of work from that of students who were told that they would be held back if their work was not good. Fear of repeating a grade appears to work only as a motivator for those students who perform adequately regardless of a threat (Kowitz & Armstron, 1961).

Still, a majority of parents, teachers, and adminstrators feel that grade retention is a sound remedy for academic failure and social immaturity. One survey of parents, teachers, and principals showed that 60% of parents, 65% of teachers, and 74% of principals favored retention when students did not meet grade level requirements (Byrnes & Yamamoto, 1984). Many believe that retaining students gives them a chance to build a foundation for future academic success, thereby increasing their chances of staying in school (Tomchin & Impara, 1992, p.200). Some students simply need that "gift of time."

Shepard and Smith (1987) found that teachers' beliefs about child development affect their philosophy on retention. Those viewing development as occurring in a series of stages were more often prone to retain students, saying there was little a teacher or parent could do if a child was unready for the next stage. Additionally, many teachers see that retainees perform much better the second year in their class, many shining as leaders. Parents also see that their child has less difficulty during the repeated year. Of course, neither parent nor teacher has any way of knowing whether the child would have done just as well if he had gone on with his peers. Nor during the year of retention, do they realize that the gains the child appears to make will disappear within the next two years. They seem not to acknowledge that while learning does take time, providing additional time does not in itself insure that learning will occur (Karweit, 1984).

Practitioners are quick to cite a multitude of reasons for the necessity of nonpromotion. Curriculum is often the culprit, disbanding the teacher's flexibility. Stepping up academic requirements in the next grade forces the earlier grade-level teachers to augment their academic requirements; kindergartners must now know how to read before they can move onto the first grade in many school districts. And certainly teachers and administrators must enforce retention when they are held responsible for maintaining specific standards. Accountability and funds are linked in most districts today and so systems like those in New York will continue to publicize their tough promotion standards, having as many as a third of their elementary students over-age yearly (Putka, 1988, p.1).

For many of these same reasons, kindergarten has suddenly become a serious matter, with its curriculum becoming much more academically oriented in the past 20 years. Over the past 30 years, the national trend has been to raise the entrance age to kindergarten; in 1958, kindergartners were required to be five years old by December 1 or January 1 but by 1985, the dominant practice was to require children to be five before October 1 to start kindergarten (Shepard & Smith, 1986, p.81). In addition to raising the entrance age, pre-kindergarten, transition rooms, and pre-first grades have been created in order to provide an extra year for unready children. Donofrio (1977) urged that these "unfavored" children be allowed to "mark time" until they are in step psychologically with their "behavioral and maturational peers" (cited in Shepard & Smith, 1986, p.84). The results of such a philosophy are vividly illustrated in this example: in 1987-1988, 22.6% of Virginia's kindergartners spent an extra year prior to 1st grade costing the state in excess of $73.5 million for that "gift of time" (Eads, 1990).

House (1988) described the current flow of the tide: "'Standards' has become a code word for retaining kids, and school districts have a much more punitive mind-set. The result is a lot more failure" (in Putka, 1988, p.1). Failure rates are at their highest in years. In Charleston, S.C., the rate is 20%, in Baltimore, 14%, in Philadelphia, 12%, and one school in Chicago attemped to fail 60% of its student population (Putka, 1988, p.1). There is no doubt that the practice of retention is prevalent among schools today, but some education watchdog groups are fighting the battle and they have plenty of ammunition to help. Research on the topic of pupil non-promotion has been clear and consistent in its findings since Keyes' 1911 study, which showed that nonpromotion was not an aid to pupil achievement or personal development.

To date, there is no consistent relationship between the achievement and ability of a student and his nonpromotion. Children with low achievement are promoted and children with higher achievement levels have been failed (Jackson, 1975). Teachers have stated that there is no purpose in "holding him back because he won't do any more if he spent 50 years in this grade," while maintaining that another student should be retained because that extra year will enable him to "catch up on those few basics that he missed." Schools with higher average achievement levels often fail larger percentages than do schools with lower average achievement levels (Caswell & Foshay, 1957). After numerous studies, Shepard and Smith (1985) declared that there simply are no set criteria for retention, an appalling consideration for a practice which leaves the pupil who is retained paying with a year of his life (Shepard and Smith, 1987). The cost is dear, and doubly so when research condemns the practice which exacts such hefty payment: "The weight of empirical evidence accrues against grade retention," the bulk of research relating that there is no positive relationship between retention and student achievement (Holmes, 1989, cited in Karweit, 1991, p.2).

WHAT HAPPENS TO THE RETAINED CHILD

What happens when a child is retained? He repeats 100% of the same course work of which he failed only 50%-60%. He sits through the same material presented in the same manner; it is a recycling process. He is placed with peers who are younger and often smaller and more immature than he. He suffers loss of status at home and among his peers. Flunking a grade is a traumatic experience, one which interferes with the learning process and incites emotional turmoil (California State Department of Education, 1991). Advocates of retention often fail to recognize, or they simply choose to ignore, the usefulness of intervention in the regular class for the at-risk child. For instance, why isn't a program developed to focus this child on the 50% of the material he did not understand? Why isn't someone aware that a change in teaching style to match his learning style might make a difference? If the first time did not result in the child meeting grade-level objectives, what makes proponents of retention believe that the act of repetition alone will achieve the desired outcomes?

PORTRAIT OF THE RETAINED CHILD

What does the retained child look like? Teachers often identify younger and smaller students as those who are most likely to be retained in addition to those who are socially immature or who are just "not catching on." Students with behavioral problems and poor attendance are often held back a year, with boys being twice as likely to be retained as are girls. A significant proportion of retentions occurs in 1st grade and kindergarten. According to research, children with the following characteristics are most likely to be retained: male, of black or Hispanic origin, from families below the poverty level, a member of a linguistic minority, handicapped, from families in which the head of the household does not have a high school diploma, and from the southeastern region of the U.S. (Illinois Fair Schools Coalition, 1985; Niklason, 1984; Rose, Medway; Cantrell &

Marus, 1983; Sater, Heaton & Allen, 1977, cited in Ostrowski, 1987).

Not coincidentally, these are the identical characteristics of children who never graduate. For years, researchers have noted the strong association between retention and failure to complete school. Godfrey (1971) determined that years of retention can be an academic handicap to students in terms of decreasing their performance, and Glasser emphasized that "once the child receives the failure label and sees himself as a failure, he will rarely succeed in school" (cited in Plummer, 1984, 9). Bachman (1971) found that two of the most significant indicators of dropping out are poor grades and grade repetition. The Illinois Fair Schools Coalition (1985) asserted that not only is there a connection between flunking and dropping out, but that nonpromotion is often responsible for forcing children out of school. The literature is teeming with recommendations for early identification and early intervention. With the bulk of retention occurring at the elementary level, the logical place for intervention would be the early grades.

RETENTION AND EQUITY

The emerging portrait of the retained child raises the question of equity. With minority and male students retained twice as often as nonminority and female students (Plummer, Lineberger, Graziano, 1986), issues of segregation and equal opportunity must be considered. Likewise, if retained children are rejected by their peers as some studies have shown (Graziano & Shaffer, 1979, Gump, 1980; Hetherington & Parks, 1979), academic and familial problems associated with retention are likely to be compounded (Plummer, 1984). Additionally, a significant proportion of students retained are routed into special education programs. A question of discrimination may be raised here. Shepard and Smith (1987) asserted that

> Retentions do nothing to promote the achievement of the affected individual or the average of the group as a whole and because the disadvantaged and minority children are most apt to be affected, retention should best be thought of as educational waste to those who most need the benefits of education. Retention has high cost and virtually no value, save the public relations advantages for the schools (p.235).

Stroup and Zirkel (1983) provided a review of the legal ramifications connected with retention practices. From the few court cases available, they determined from their look at cases that retention policies should use multiple criteria, avoid radical changes, and not disproportionately affect any single minority group. According to Walden and Gamble (1985), legal challenges to school district retention policies are increasing. With the staggering amount of research showing retention either to be of no benefit or actually to be deleterious, the onus of proving this treatment to be in the best educational interest of the child may prove to be a very difficult one for the school districts employing this practice.

ALTERNATIVES TO RETENTION

Critics of retention are certainly not arguing that students be allowed to slip through the system without learning. They are saying, and emphatically so, that nonpromotion is not the answer to America's dilemma of what to do with students who are identified as unready to pass on to the next grade. A host of alternatives to repetition of grade is available. The literature, brimming with studies condemning retention, is likewise replete with suggestions and programs designed to address the needs of the child who is socially immature or who is behind academically. The majority of these suggestions have grown from a philosophy that "an ounce of prevention is worth a pound of cure."

While retention may have been the sole solution to problems created by the advent of graded schools, times have changed, bringing forth a deluge of ideas that were not available at the beginning of this century. One obvious alternative to retention is remediation as described by Hess (1978): "The problem with graded repetition lies not in the repetition, but in carrying out the practice on a unitary, all-or-nothing basis. The best alternative to repetition is repetition, but within the context of existing grade levels" (p.162). Promotion with remediation may be accomplished in a number of ways: peer and cross-age tutoring, summer school programs, special help sessions on weekends and during vacation periods, and after-school programs. These approaches to remediation allow for more time-on-task during the specific session and do not cause students to miss regular classroom instruction. Year-round schools also address the need for immediate remediation. Cooperative learning strategies and learning styles inventories also prove to be useful for the at-risk child. A helpful factor in promoting these strategies is cost; compare $800 for remediation to $3000 or more for retention per student yearly (1991 dollars; Illinois Board of Education, 1990, cited in Reynolds, 1992).

Nearly at the opposite end of the spectrum are considerations of politics, research, and accountability. Shepard and Smith (1989) have advocated action research to be conducted by teachers and political activity to inform legislatures and school boards about the effects of retention and to work for policy changes that would institute viable options to retention. Norton (1990) agreed that the scope and concept of accountability should be broadened to involve a look at what goes on in the classroom through portfolios of students' work and student interviews. Rose, Medway, Cantrell, and Marus (1983) have called for a focus on classroom instruction, and Taylor (1978) emphasized the teacher's role:

> "In the final analysis it is the classroom teacher who has the greatest knowledge of the student's achievments, hence the teacher should be the first person concerned with identifying problems and correcting them. This is part of the teacher's job" (p.25).

Identifying the at-risk child's specific problems as early as possible and then prescribing an individualized plan of instruction serves to increase the child's success. Providing for differing progress and individual needs by adapting expectations to a more realistic level can be instrumental as well. Outcome-based education, nongraded elementary schools, and multi-age grouping promote these strategies. Implementing more developmentally appropriate curriculum and tossing out readiness testing as a tool in deciding grade placement are essentials for lowering the retention rates. Parental involvement and parental assistance programs have been found to be helpful as well.

Rather than developing new programs, Byrnes and Yamamoto (1986), among others, propose smaller classes with more individualized instruction and increased remedial instruction opportunities. Smaller classes serve not only as an alternative to retention but also as a preventive measure. The 1988 Phi Delta Kappa/Gallup Poll showed that 77% of parents believed that having a small class made "a great deal of difference" in student achievement. Teachers also profess that smaller classes have a positive effect on student achievement and development due to an increase in morale and the fact that the teacher has more time to interact with the student. Shapson (1980) found that small classes made a large difference to teachers in terms of attitude and expectations; Project STAR noted that teachers with smaller classes reported fewer problems related to time, and Whittington et al. (1985) stated that teachers with smaller classes related better classroom climate, improved student behavior, more time for planning and preparing for instruction, smoother noninstructional times, and better teacher/pupil interactions (in Achilles & Moore, 1986).

While educators have debated the issue of class size for years, conclusions are still controversial. Glass's (1982) meta-analysis concluded that class-size reduction did have a substantial effect on pupil achievement. Ryan and Greenfield (1975) hypothesized that greater individualization of instruction was most likely the operating factor in increasing student achievement in smaller classes. Studies (Bourke, 1986; Johnston, 1990) seeking to identify factors in smaller classes that would contribute to increased achievement listed these: increased interaction between teacher and pupils, more individualization of instruction, better diagnosis of each child's learning needs, possibilities for more active involvement of students in learning tasks, and less time spent on classroom management (Folger, 1989). These processes are the very ones that are likely to benefit students at risk of school failure (Slavin & Madden, 1987). Tennessee's STAR offered conclusive support that smaller classes did indeed have an advantage over larger classes, particularly in reading and mathematics. Achilles and Moore (1986) concluded from STAR that class size is a facilitative variable, agreeing with other researchers that smaller classes may permit teachers to spend more interactive time with students, etc. The literature review points toward smaller classes as a strong alternative to remediation, perhaps the best available considering its preventive tendencies.

REVIEW OF THE RESEARCH

Research on the effects of grade retention began as early as 1909 and continues today. The preponderance of these studies occurred during 1960 and 1975 and was devoted to the effects of nonpromotion on student achievement and on the social and emotional development of these children. The theoretical and empirical support for retention is sparse and questionable. The two largest reviews of the research by Jackson (1975) and Matthews and Holmes (1984) concluded that the practice of retention has either no effect or a negative effect on academic achievement and the social and emotional well-being of those children not permitted to progress with their peers.

The studies reviewed in this section are a sampling of those which examine the effects of retention on children in the early primary grades and consider how retention affects academic and/or social outcomes. The majority of these studies failed to look at the influence of retention over long periods of time, and few examined the actual retention model employed (Ostrowski, 1987). Research that found positive effects on academic and social outcomes are reviewed first, followed by research finding either no effect or negative effects on these outcomes.

STUDIES THAT FIND POSITIVE ACADEMIC EFFECTS OF RETENTION

Proponents of nonpromotion contend that promoting students who have not acquired the grade level objectives will only result in further frustration for the children. They contend that giving some students the "gift of time" is necessary since children learn at different rates. Early research by Buckingham (1926), Keyes (1911), and McKinney (1928) indicated that nearly one-third of the children studied displayed favorable academic gains during the retained year. These studies did not look at long-term effects. Lobdell (1954) stated that approximately 69% of retainees might be expected to evidence good or fair progress through careful selection. Stringer (1960) also found that retention seemed to be more helpful than harmful to students' academic performance, but he did note that gains diminished the second year, which is often the case. However, an extensive two-year study conducted by Sandoval and Hughes (1981) using 146 first graders identified as potential repeaters did find that the successful retained group was inferior to the promoted group only in mathematical achievement (Plummer, 1984). A cautionary note was added to this research: the study evaluated retainees for only one year after nonpromotion.

Additional support for nonpromotion comes from Kerzner's (1982) research using 56 students who had progressed and completed one grade beyond retention. He found that retained children in second and third grade exhibited the greatest positive effects. A year later, Vollrath compared a group of retained K-3 pupils with those who were recommended for retention but were promoted.

Matching was done on a cognitive abilities test. Same-grade comparisons were conducted at grade 3 and 6. He found that the retained group was significantly higher than the promoted group.

Some of the strongest evidence in favor of retention comes from an innovative program in the Greensville County school system in Virginia. The system abolished the existing social promotion policy and replaced it with a policy enforcing an academic mastery program implemented by Owens and Ranick (1977). The program has produced respectable success rates; students previously scoring in the bottom 20% to 30% nationwide on achievement tests have risen to the top 50% to 60%. The number of retainees and dropouts has fallen also and the community, students, and teachers are satisfied with the policy. Yet, this is not the typical retention-by-recycling program. Interventions are applied throughout the system in the form of designing instructional programs specifically to meet the needs of slow learning students, allowing partial promotions, and implementing block scheduling to permit more individual contact between teacher and student.

STUDIES THAT FIND POSITIVE SOCIAL EFFECTS OF RETENTION

The second major factor in retaining a child is that of social maturity. Goodlad (1954) found less damage to a child's social relations with peers among first grade retainees than for those retained in later grades. This belief has influenced retention practices greatly, accounting for the fact that the highest incidence of nonpromotion occurs in the first grade (Peyton, 1968; Rose, Medway, Cantrell, & Marus, 1983). The bulk of research favoring retention functions from this supposition that students will be less adversely affected the younger they are retained.

In 1976, Horn surveyed primary teachers in a school operating from a nativist perspective, that is, students were retained because of social immaturity. He also interviewed students who were retained. Teachers reported that retained students developed better self-concepts through their successful second year. Students likewise said that they felt good about themselves and school. This study suggested that teacher attitude is important in determining their emotional impact on students, but it must be noted that this was not a controlled study and that the methodology relied on opinion.

Chansky (1964), Chase (1968), Finlayson (1977), and Ames (1981) conducted research concluding that retention had no effect on emotional well-being or that its effect is positive. Chansky ascertained that there was no difference in personality attributes between promoted and retained children as measured by the California Test of Personality. Chase indicated that 75% of the 65 first-, second-, and third-graders studied had no emotional upset. In studies by Chase and Ames, teachers and parents were supportive of retention, stating that they had observed positive changes in their children. Plummer (1982) and Jawarski (1985) conducted research also indicating that parents and teachers sensed an increase in the retainees' self-concepts. Plummer discovered that students who were not promoted actually had higher self-concepts than promoted pupils. Finlayson studied from 1973-1975 the self-concepts of first-graders divided into nonpromoted, borderline (promoted), and promoted. He found the nonpromoted students continued to increase their self-concepts during the second year. Teachers saw 96% of the nonpromoted children as having the same degree of or more positive self-concept. Parents also viewed the experience as positive for their children. What Finlayson neglected to remark was that the borderline group of students demonstrated the highest final score on the measurement of self-concept. Does this finding imply that promotion for potential retainees improves their self-concept (Ostrowski, 1987)?

STUDIES THAT FIND NEGATIVE EFFECTS ON SOCIAL OUTCOMES

Research looking specifically at the effect of retention on social outcomes is neither abundant nor particularly emphatic on either side of the coin. There do exist, however, at least as many studies declaring retention to have negative effects on a child's self-concept as concluding that the effects are either positive or indeterminable. The problem in determining effect centers around the question of which comes first: does poor self-concept initiate failure or does repetitive failure prompt low self-esteem?

In 1944, Sandin designed a study in which he addressed the social and emotional adjustment of promoted and nonpromoted students. He found that retainees generally isolated themselves from their regularly promoted counterparts and did not consider their classmates appropriate companions. The resulting lack of an appropriate social environment contributed to their feelings of discouragement and intent to quit school. Conclusions from Caswell and Foshay (1957) agreed with Sandin, stating that nonpromoted students suffered from depression and discouragement. Johnson's (1981) research concluded that children experiencing chronic failure eventually develop feelings of helplessness. These children are subjugated by criticisms from classmates, parents, and neighbors.

Interviews with retained children have shown that they perceive flunking a grade as punishment and a stigma (Byrnes & Yamamoto, 1985). They associated retention with "sad" and "bad" feelings, feelings of being upset. They rated nonpromotion as being worse than most events in their life; death of a parent was one of the few incidents classed more terrible than failing. Shepard and Smith (1985) found that 84% of the retained children in their matched groups also associated "sad" feelings with nonpromotion.

Other researchers have noted that retention negatively affected the social and affective development of students. Goodlad (1954) evaluated 73 nonpromoted students from six elementary schools and 150 children from five elementary school who were borderline but were promoted. Using the Kuhlmann-Anderson Tests, the Metropolitan Achievement Tests, the California Test of Personality, and the Haggerty-Olson-Wickman Behavior Rating Schedules, he found that retained pupils were less accepted as friends and that the pupils' adjustment scores dropped following the retention, dipping from 4.94 to 4.71 while the promoted group increased its mean score from 4.42 to 5.00. Morrison and Perry's (1956) research supported Goodlad's findings. They stated that over-age children were generally least accepted in the classroom. Godfrey (1972) conducted research in North Carolina with 1200 students, confirming that retention has a detrimental effect on student self-concept and attitude. Using the Tennessee Self Concept Scale, Godrey found that retained students scored lower on every sub-scale than promoted students and were inclined "to doubt their own self-worth, have little confidence in themselves, see themselves as inadequate in social and family situations, and have an unfavorable view of their own behavior and moral worth" (cited in Ostrowski, 1987, 20). She suggests the need for alternatives which meet the educational and psychological needs of the children.

STUDIES THAT FIND NO EFFECT OR NEGATIVE ACADEMIC EFFECTS OF RETENTION

Research finding negative or no effect on academic achievment from retention in the primary grades is plentiful. As far back as 1911, Keyes conducted a four-year study with 5,000 students in an urban school district, producing results indicating that of the large number of retainees, 20% did better academically, 39% showed no change, and 40% actually did worse (cited in Plummer, 1984). There are several other supporting studies from the early part of the century. Buckingham (1926) found that only about one-third of several thousand children did better academic work after nonpromotion. Klene and Branson (1929) determined that potential repeaters profited more from promotion than did repeaters from nonpromotion. Children were matched on the basis of chronological age, mental age, and sex, with half the students promoted while the other half were retained. This was one of the few studies using an experimental design. Arthur (1936) also concluded from a group of 60 first-grade repeaters that the average repeater did not learn more in two years than the average nonrepeater learned in one year. Again, the students were matched by mental age as based on intelligence testing though an experimental design was not used.

Farley, Frey, and Garland (1933) continued the effort to distinguish between repeaters and nonrepeaters. They compared students with low IQ's who had repeated several grades with children of the same ability but who had not been retained. Those children who had repeated grades were found not to have done as well in their school work as those who were promoted. These researchers indicated that, in this case, retention was likely to be more of a deterrent than a catalyst to acceptable academic standards. Goffield's 1954 results supported these findings.

Subsequent research adds backing to these early conclusions. Matching on sex, race, age, socioeconomic level, mental ability, reading achievement, and type of classroom assignment, Dobbs and Neville (1967) found that retention did not improve reading or mathematics scores in their study of 30 pairs of children. Abidin, Gollady, and Howerton (1971) concurred with the findings of Dobbs and Neville, offering further support for the continuous decline in both achievement and ability level as a function of retention. Their research involved a group of 85 students who were retained in first and second grade and whose achievement and ability declined through the sixth grade.

A year later, Godfrey (1972), in a retrospective study of North Carolina students, concluded that not only does nonpromotion not enable students to catch up academically, but that it is actually an academic handicap. His population consisted of 1200 sixth-and seventh-graders. Nonretainees were reading at a 6.8 grade level while students who had repeated one grade were reading at a 5.2 level, and those students repeating two or more grades had dropped to a 4.5 grade level. In mathematics, nonretainees averaged in the 27th percentile, one-time repeaters in the 10th percentile, and those who had repeated twice or more were in the 5th percentile.

In a comprehensive review of the research literature on the effects of grade retention, Jackson (1975) found a bias in the design of a majority of the studies. He identified three design types. Type I studies compared nonpromoted students with promoted students. Jackson asserted that a bias in favor of grade promotion occurred because students with difficulties were compared with students who normally did not have such difficulties. Type II studies compared retained students before and after their retention. With no comparison group, there were no controls for improvement due to factors other than retention. This design failed to assess the advantages of grade retention. Type III designs involved an experimental design in which students recommended for retention were randomly promoted or retained, and were compared at a later time. This was the only bias-free design. Unfortunately, only three of the original 44 studies were of this design type. Random assignment of children could raise ethical questions which accounts for the rare use of this format. Jackson called for much more high quality research, stating that the available research was too poor to make valid inferences concerning the benefits of retention.

During the 1980's, research did get stronger. North Carolina's Department of Public Instruction followed with a major study in 1983 that compared pairs of promoted and nonpromoted first graders to determine how they had performed in subsequent years. These students had identical reading achievement test scores in grade one, but three years later the promoted students had higher scores, gaining each year. While the nonpromoted students did have a higher class rank at the end of the first year, the difference had become insignificant by the end of the third. The weight of this study stems from its large sample size and the fact that it was longitudinal. Nichalson's 1984 study investigating students recommended for retention in two Utah school districts confirmed the findings in North Carolina. Peterson (1985) also found that any increased achievement from the retention experience diminished within two years.

In 1983, Holmes examined eight reports of studies in which retained students had been matched on the basis of achievement test scores with promoted students. Standardized achievement test scores were used as measurements of the dependent variable achievement. He concluded from his analysis that retained students fall behind during the year in which they are retained and are never able to recover.

The following year, in conjunction with Matthews, Holmes conducted a meta-analysis of 44 retention studies. Specific qualifications for selecting studies were established. The study presented the results of original research of the effects on pupils of retention in the elementary or junior high school grades, contained sufficient data to allow for the calculation or estimation of an effect size, and compared a group of retained pupils with a group of promoted pupils (Holmes & Matthews, 1984, p.228). The effect of retention was measured in 31 studies, with findings indicating that the promoted group had achieved .44 standard deviation units higher than the retained group. Twenty-one studies yielded these results on personal adjustment: the retained students scored .27 standard deviation units 9<27 effect size) below that of promoted students. On measures of self-concept from nine studies, promoted pupils outscored the retained pupils by .19 standard deviation units. Conclusive evidence was found to illustrate that the potential for negative effects from retention far outweighed positive outcomes.

Other researchers have offered support from their studies of long-term effects. In 1986, McDaniel found that over a five year period, retained students achieved significantly lower on mean NCE scores than their promoted counterparts. A case study done by Routh in 1986 looked at achievement data for three sets of students at the elementary level; she found that retained students showed very little improvement and often regressed on academic achievement tests. A study by Peterson, DeGracie, and Ayabe (1987) examined longterm effects of retention on the academic achievement of a group of elementary-age students. Retained students received remediation through individualized plans while the comparison group was socially promoted but did not receive remediation. Retained students performed better on the California Achievement Test during the first year, with most of the retained students losing their gain by the second year, and there being no difference in scores by the third year. Baenen (1988) followed 243 matched pairs in a five-year study. Students had received remedial services in this progam. Results showed that retained students performed better the first three years, but that the positive effects faded and became negative after that. The average effect size across all five years was .15.

In 1989, Holmes meta-analyzed 63 studies of the effects of retention in elementary school on academic achievement and non-academic outcomes. Fifty-four of the studies reported negative effects from grade retention for an overall mean effect size (ES) of -.26 standard deviation units across grade levels and outcomes. Academic achievement was found to have been affected the most adversely (ES = -.31), with attendance having an effect size of -.23; personal adjustment, -.21, and attitudes toward school, -.18. The results remained consistent when retained and comparison-group students were matched on prior achievement, IQ, sex, socio-economic status, and grades. Additionally, these findings were constant with earlier research analyses.

The available research on retention indicated that nonpromotion is not likely to improve the academic achievement of children. Many studies questioned the validity of the practice of retention practices, while most identified it as having negative effects on achievement.

Results on the social and emotional effects of nonpromotion were mixed. Overall, retention was labeled as a crude intervention (Sandoval, 1984) and suggestions for alternatives were emphasized.

REVIEW OF CLASS SIZE RESEARCH

Class size research conducted prior to 1920 was largely concerned with the effects of large classes on grade-to-grade promotion rates (Cornman, 1909; Boyer, 1914; Bachman, 1913; Elliot, 1914; Harlan, 1915 cited in Mitchell, Carson, Badarak, 1989). Rice (1909) conducted the first empirical research in this category. He concluded that there was no relationship between class size and student achievement; no statistical data were included. Numerous studies of this sort followed with similar results. It was not until research designs were improved that results began to show that class size did affect student achievement. Fully randomized research designs were not employed until 1930.

Davis and Goldizen (1930) concluded from their study of 140 seventh grade history students of medium ability that students in the large class were at no disadvantage. Students had been divided into three classes: 70 in one class and 35 in each of the other two. Whitney (1930) and Willey (1932) found contrary results in their studies involving elementary school students in Colorado. They used

12 groups of 20 students matched against 12 groups of 40 students for regular classroom instruction. Achievement of students in both groups was measured and compared before and after instruction. Eighty percent of the comparisons favored smaller classes.

The events of World War II put research in class size on the back burner for awhile, but renewed interest arose with the influx of baby boomers into the schools. In a carefully designed study by Balow (1969), a program which reduced class size from 30 to 15 for reading instruction was implemented. A stratified random sample of seven elementary schools with all children in grades one through three participated in the program. Measures of reading achievement were analyzed yearly. Cumulative results were compiled at the end of three years, and Balow found that students in the experiemental program scored significantly higher than other children at the end of each year, with influence of the program being cumulative.

In 1978, using the newly developed technique of meta-analysis, Glass and Smith conducted a review of the class size literature. Identified were 77 studies which compared larger and smaller classes, totaling 725 comparisons of pupil achievement in classes of at least two different sizes. Research spanned 70 years and was divided into subgroups based on age, grade, and length of time students were exposed to instruction within a specific class. They found small gains in the 63 "poorly controlled" studies and large gains in the 14 "well controlled" studies. They concluded:

> The relationship is seen most clearly in well-controlled studies in which pupils were randomly assigned to classes of different sizes. Taking all findings of this meta-analysis into account, it is safe to say that between class-sizes of 40 pupils and one pupil lie more than 30 percentile ranks of achievement. The difference in achievement resulting from instruction in groups of 20 pupils and groups of 10 can be larger than 10 percentile ranks in the central regions of the distribution. There is little doubt that, other things equal, more is learned in smaller classes (cited in Mitchell, Carson, Badarak, 1989, p.29).

The Educational Research Service (ERS) fired a round of criticism at Glass and Smith's conclusions as well as the technique of meta-analysis itself. In 1978, the ERS conducted its own review of about 80 studies, sorting them by results favoring small classes, large classes, and studies showing no difference. Most of these studies were correlational focusing on differences in the range of class size between 25 and 30. The ERS reviewers concluded that differences were usually nonsignificant, positive or negative. Slavin (in press) explained that such findings were due to the range of class size studied and the fact that no consideration was given to study characteristic or quality.

Glass, Smith, Cahen, and Filby reanalyzed class size research in 1982 and the ERS conducted an updated review of the literature in 1986. The reanalysis by Glass et al. confirmed their original findings that there was a statistically significant negative correlation between student achievement and class size. The ERS review concluded that effects of class size were relatively consistent in grades K-3, slight in grades 4-8, and essentially nonexistent in grades 9-12.

Slavin (1984, 1989) criticized both conclusions from Glass et al. and the ERS review, stating that neither had considered the quality of the critical evidence. Using an abbreviated form of a review technique called best-evidence synthesis, Slavin analyzed eight studies which he determined to deal most directly with the question of optimum class size in the elementary schools. For each study, effect sizes were computed. Results showed that substantial reductions in class size generally had a positive effect on student achievement; the median effect size was +.13.

Policy makers in Tennessee wanted the question of class size and achievement answered once and for all. In 1985, a cooperative project that would endure for four years involving the State Department of Education, a four-university consortium, and 42 local school systems was begun. The legislature required that the project select schools from inner city, suburban, rural, and urban areas. Seventy-nine schools with students randomly assigned to small (13-17 students), regular (21-25 students), or regular with an aide (21-25 students) classes participated in STAR. There were 101 regular classes, 99 regular with an aide classes, and 128 small classes. Teachers were randomly assigned to one of the three class types while students were initially randomly assigned to a class type and remained with that class type throughout the study. New students were randomly assigned in accordance with vacancies. Students were tested yearly on the Stanford test in K, 1, 2, and 3 and on the state-developed criterion test in grades 1-3. The pupil was the primary unit of data collection, and the class was the unit of analysis.

Results from STAR were conclusive. Pupils in small classes made significantly greater gains than other pupils. The class size effect was found equally in all four locations and favored the S condition in all four grade levels with the greatest gains visible in K-1 (Word et al., 1990; Nye, Achilles, Zaharias, Fulton, & Wallenhorst, 1992). The effect sizes were about one-fourth of a standard deviation among students and ranged from about one-third to two-thirds of a standard deviation among class means. Students in grades 2-3 continued to benefit from small classes though the gain was not as large as in the previous grades. Achilles and Moore (1986) contributed these gains in achievement to the facilitative nature of small classes.

While few scholars debate the issue of whether small classes contribute to increased student achievement, the question of cost effectiveness has arisen. Use of smaller classes is expensive; yet, as Slavin (1992) stated: "Even very expensive early interventons can be justified on cost-effectiveness grounds alone if they reduce the need for later and continuing remedial and special education services, retentions, and other costs" (p.12). Early intervention, even when expensive, is receiving more widespread acceptance because it has been determined that it pays back its costs. The trend is visible in the states that have either acted to reduce class size--Arkansas, Indiana, Florida, Nevada, South Carolina, Tennessee, Texas, West Virginia, Alabama, and Pennsylvania--or are considering class size reduction--California, Minnesota, New Jersey, Virginia, and Wisconsin. Studies of the magnitude of STAR have enabled policy makers to comprehend the significance of smaller classes on children's learning.

WHAT TO EXPECT

According to the literature, the retained child is a male of minority race, usually black or Hispanic. His family is below poverty level, with the head of household failing to hold a high school diploma. The child lives in the southeastern United States and attends an inner-city school.

These same characteristics are expected to be prevalent in the retained child of the STAR database. The student is likely to be a black male from a low income family who attends an inner-city school in Tennessee. No data from STAR are available to determine whether the child's parents hold a high school diploma. Whether the retainee is in kindergarten or first grade, his portrait will be very similar if we believe prior research.

CHAPTER III

METHODOLOGY

PRESENTATION OF STUDY

This study has examined the effects of class size on factors related to retention as evidenced by pupil achievement data from Project STAR. Given its detail, size, and longitudinal nature, the Project STAR database provided additional insights into issues surrounding pupil retention in grades. Chapter One detailed the problems of retention and the need for alternatives. The rationale for examining the rate of retention as affected by class size was given. The following questions were asked: 1)what picture of the retained kindergartner and retained first grader is given by demographics, and 2)if a retained student is subsequently placed in either a small class, regular class, or regular class with an assistant, what are the differences in achievement for retained students based on class-size placement?

Chapter Two offered a review of the literature researching grade retention dating back to the beginning of the 1900's. Retention has been a common practice for nearly a century despite cumulative evidence relating no benefit or negative effects. Today, when the public demands accountability and the government chooses to link funding with achievement, educators must curtail practices which deplete resources of mind and money. This study adds further information about retention, such as whether small classes will prove advantageous in this effort.

PROJECT STAR

The extant data base of Project STAR provided the data for the current study of retention and class size. No new data were collected; STAR data were reanalyzed for purposes of the present study. Cooley and Bickel suggested that decision-oriented research make use of already existing data. Policy making too often depends on opinion of the policy makers rather than on information produced by research. The current study followed Cooley and Bickel's recommendation. A brief description of the STAR database and processes will be helpful in explaining the methodology of this study.

STAR used a within-school design and random assignment of teachers and students to the three class conditions of small, regular, and regular with an aide. This in-school design reduced the major sources of possible variation in student achievement attributable to school effects.

Initial selection of participating schools was made with the choice of schools within systems determined partly by school size. The in-school design required that enrollment be large enough to provide at least one class type at each grade. Grade-level enrollment determined the number of classes of each type within each school. The 79 elementary schools selected provided approximately 100 classes of each type. These schools served rural, urban, suburban, and inner-city students with approximately 6,500 students participating in Project STAR in kindergarten. In 1985-1986, there were 128 small classes, 101 regular classes, and 99 regular classes with aides. Students in small class in kindergarten remained in small class through grade three. There were approximately 7100 first graders. All students entering Project STAR after the initial year were placed in class type randomly. Attrition of students and schools was accounted for by oversampling.

STAR was a randomized experiment employing the control-group design of Campbell and Stanley (1963), Design Number 6. This design uses post-test analysis only. Project STAR's primary analysis consisted of a cross-sectional analysis of data from all students participating in project classes at each grade level. In addition, longitudinal analyses were conducted in which data were analyzed for students who were in the project in the same class type for consecutive years. Analyses-of-variance were utilized. Appendix A contains basic information on the STAR design.

Project personnel collected data about student achievement, development and variables, other than class size, that might have affected achievement. Data collection instruments included the Stanford Achievement Test (SAT), Tennessee's Basic Skills First Test (BSF), the Self-Concept and Motivation Inventory (SCAMIN), school and system profile, principal profile, teacher profile, teacher log, grouping questionnaire, parent/volunteer/teacher interaction questionnaire, teacher problem checklist, teacher exit interview, aide profile, aide questionnaire, aide log, roster, and special programs form. Yearly, data from the measurement instruments were analyzed in subsets: the SAT achievement scales, the BSF performance tests, and the SCAMIN. Multivariate test statistics were used for each subset.

SAMPLE

The STAR database was used as a means to analyze the phenomenon of retention and class size. The population for this study is the students who were retained at the end of kindergarten (1984-85) and those who were retained at the end of grade one (1985-86) in Project STAR. STAR began in 1985 with students who entered kindergarten during that year. Entry profiles of students showed whether a student had been retained in kindergarten (1984-85). Student records related that 253 youngsters had been retained in K (1984-85) and entered STAR in K (1985-86). Students who entered the STAR database in grade one in 1986 had been held back in first grade or were new to the project. Over-age students in K (1985) were either a)kept out of school for some reason or b)retained in grade in K. Kindergarten was not required in the state of Tennessee in 1984-85 and so some students entered school for the first time in grade one.

Students who entered STAR for the first time and were six years nine months and twenty-two days (6.8years) and younger as of October 1, 1986 were considered new first graders. Those students who were approximately six years eleven months (6.9 years) and older at this time were considered to have been retained. Students who had been retained in kindergarten were identified by teachers who marked such information on student forms; this information was then added to their record on the STAR database.

The STAR database followed students from kindergarten through third grade. If a student in the STAR cohort left or was retained, a new student was added by random replacement to the cohort. No additional data were collected for the student who left the STAR cohort. In order to determine the effects on retained students, retained students were identified from student records and/or picked up new students who entered STAR each year and who were approximately one year older than their "regular" age mates. For example, in 1986-87 (grade one) 2276 new students entered STAR; 1152 of these were "overage," defined as at least 6.9 years as of October 1, 1986. An age of 6.9 years is approximately equivalent to six years, eleven months.

Entry age of students into kindergarten is determined by the State Board of Education. In Tennessee, a child may enter kindergarten if he is no less than five years old on or before September 30. A child enrolling in first grade must be no younger than six years old on or before September 30 of the enrollment year. He must enter kindergarten or grade one no later than his seventh birthday. Kindergarten was not required at the time of STAR in Tennessee.

SAMPLE VERIFICATION

Teachers identified 253 kindergartners as having been retained in 1984-85. These youngsters entered STAR in 1985-86 as repeating kindergartners. At this time, 6041 first time kindergarten students entered STAR. A frequency distribution of the 253 retainees related that 11 (4%) were 5.8 years or younger; 242 or 96% of this group were 5.9 years or older as of October 1, 1985. The mean age of new enrollees was 5.4 years while the mean age of retained kindergartners as of October 1, 1985, was 6.2 years. These students would then be at least 6.9 years (approximately 6 years, 11 months) when they entered first grade, the age selected as an indicator of retention for the grade one sample. Confidence in selecting this age as an indicator of retention was established with such a high percentage of retained kindergartners showing at least 5.9 years for kindergarten entrance in September, 1985, and subsequently would be 6.9 years for grade one in September, 1986.

Grade one retainees were identified as overage students. Grade-one children 6.9 years and older (as of October 1, 1986) were determined to have been retained using two methods. One, an age frequency distribution of the known retained kindergartners was run. The mean age as of October 1, 1985, was calculated. Two, as a follow up to this analysis, a pilot study was prepared using a sample of STAR students from the Knox County , Tennessee schools. This system offered both urban and rural schools that were racially mixed, thus, being representative of the original population.

A list of overage first grade students from 1986 was then compared to the 1985 list of students who had originally started in Project STAR. In May 1993, the compiled list of 63 older students was then checked by Knox County's coordinator of research and evaluation. He identified 50 students who were still enrolled in 11 middle schools in his district. He then forwarded to these schools a letter indicating the need for verification of retention. After checking permanent school records, guidance counselors at the schools responded to the question of whether specific students had repeated kindergarten or first grade or had spent a year in a transition class by checking the appropriate place by the student's name. It was necessary to use student names because student identification had been changed to social security number with the inception of Project STAR.

A coordinator of research and evaluation verified that 50 children from the list of 63 older Knox County students were still enrolled in the district. Of these 50 students, guidance counselors confirmed that another five students had left the district. Of the remaining 45 students, 10 (22%) were identified as having no record of retention while 33 (78%) were known to have been retained. A 7.1% attrition rate of students was found.

Using the age of 6.9 as of October 1, 1986, 1152 students of the 2277 students new to STAR as first graders in 1986-87 were identified at the end of grade one as having been retained while 1124 were recognized as new students entering STAR. (One student had no test record.) In Tennessee, 5-6% of children in K-3 are retained each year, with more than twice the number retained in first grade as in kindergarten or ingrades two or three. In 1986, the retention rate in the state of Tennessee was 9.2% in grade one (Record of Pubil Progress in Public Schools--1986-1987).

DATA COLLECTION

The Center of Excellence for Research in Basic Skills extracted data from the STAR database for the population of those students retained in kindergarten and grade one as requested, tracking them through grade three. The mean and standard deviation of the scores for the total reading and total math sections of the Stanford Achievement Test (SAT) were collected for both students not retained and those retained (younger and older) students by class type at the end of kindergarten and grades one, two, and three. Total percent passing was calculated for these same parameters on the BSF. (BSF is not given in K.) Total number of students tested was also given for each section of the test, disaggregated by class type within not retained and retained (younger and older) students. The number of students whose scores were available for the two sections of the SAT and BSF were not constant for all test times. This minor variation occurred because all students were not always present for each of the three parts of the test. The difference in number of students tested can be assumed to be reasonably equivalent among class types due to the randomness of student placement.

This information was collected on students who began school in kindergarten and those who entered in grade one and who may have begun their education without kindergarten. Demographics of sex, race, socio-economic status (determined by free and non-free lunch), class size distribution, and school type distribution were collected on students at the end of kindergarten and grade one by class type and not retained and retained (younger and older) sub-sets. Again, numbers vary from sub-set to sub-set due to incomplete data on students, or due to student attrition.

ANALYSES

This study used post-test analysis of the students' results on the SESAT II test at the end of kindergarten, and the results on the SAT at the ends of first, second, and third grades, and on the BSF test at the end of grades one through three. An analysis of variance (ANOVA) was computed on scores for small (S), regular (R), and regular with an aide (RA) classes for retained kindergarten students and retained first graders as well as those who had not been retained. Computer analysis provided F ratios and F probabilities. Trends were identified by comparing those students who had been retained to those who had not been retained. The frequency and percent of placement by class size and school type were also calculated. Chi-square was used to calculate significance at the .050 level for demographics of retained and not retained students at the $p \leq .05$.

CHAPTER IV

PRESENTATION OF DATA

INTRODUCTION

Chapter Four provides the findings from the main study Results are reported from the pilot study as they help establish the confidence level by which overage population for the sample of retainees was taken. Results are then reported for retained kindergartners and kindergartners who were not retained. Results are also given for the retained first grade students and those not retained. Both sets of students are tracked through grade three. Trends are identified between retained and non-retained groups. Data are summarized in table format.

DEMOGRAPHICS OF RETAINEES ENTERING KINDERGARTEN AND FIRST GRADE

The retention rate in kindergarten in 1986 in Tennessee was 3.7% (Record of Pubil Progress in Public Schools--1986-1987) while 4.0% of the kindergartners were retained in Project STAR. Distribution of retained students was approximately equivalent to that of new kindergarten students with no significant difference as calculated by chi-square (.32 with two degrees of freedom). Table 1 shows that previously retained students (n = 253) from kindergarten (1984-85) were placed in STAR kindergarten in 1985 in S at 28.1%, in R at 39.1%, and in RA at 32.8%. New kindergartners were distributed as follows: in S at 30.1%, in R at 34.5%, and in RA at 35.4%. Placement of retained and not retained students in STAR in kindergarten by class type is summarized in Table 1.

Distribution of first graders retained and then entering STAR in 1986-87 varied from that of the retained kindergartners; a lower percentage of first graders was assigned to small classes in grade one than that in kindergarten. Chi-square was significant. Table 2 shows that previously retained students (n = 1152) from grade one (1985-86) were placed in STAR grade one in 1986 in S at 14.4%, in R at 45.3%, and in RA at 40.3%. New first graders (n = 1124) were placed in S at 19.0%, in R at 43.5%, in RA at 37.5%. Though students were randomly assigned to class size, there were fewer students in S proportionately than the other two conditions because the (S) condition could not exceed n = 17 by initial placement. More than two and a half times that number were in the R condition and more than twice that were in RA. Placement by class type is summarized in Table 2.

As has been evidenced in previous retention studies, the rate of retention is higher among boys than among girls. The same is substantiated. (See Table 3.) Slightly more than two times as many males were retained in kindergarten as were females: of the retained population, 69.2% were males and 30.8% were females. The population consisted of 51.4% males and 48.6% females. Of the retained population, approximately 69% were male and 31% were female. This difference is significant.

The pattern of retention for first grade males and females parallels that of the kindergarten retention rate though the difference is not as great. Of 1153 retaines, 714 or 62% were males and 438 or 38% were females. Approximately twice as many boys as girls were retained. Chi-square was significant. Again, this information coincides with findings from previous studies of this nature.

In comparing the male-female retention pattern of first graders and kindergartners, note that 69% of the kindergarten retainees were male and 62% of grade-one retainees were male, a difference of 7%.

Disaggregation by race produced unexpected results at the kindergarten level. Of the white population of 4216, 4.8% entered STAR as kindergarten retainees from 1984-85. Of the minority population of 2078, 2.5% entered STAR as kindergarten retainees from 1984-85. The population of not-retained students entering STAR in kindergarten (1985-86) consisted of 66.5% white and 33.5% minority while the population of previously retained (1984-85) pupils entering STAR in kindergarten in 1985-86 was 79.4% white students and 20.6% minority students. These statistics do not coincide with the popular belief that the majority of retained students are minority children. The difference as measured by chi-square is significant. Table 4 summarizes these data.

Analysis of retention by race in grade one showed no significant difference. Sixty percent of the new entrants in grade one were white with 51.5% retained. Of the minority new entrants in grade one (n = 910), 49.3% entered as retainees. Of the pupils retained (n = 1151), 61% were white and 39% minority, while of the non-retained entrants 59% were white and 41% were minority. See Table 4. Retention among kindergartners showed more than twice as many white students were retained as were minority children; grade one showed an almost equal number of retentions between the races.

Breakdown in kindergarten by socio-economic status was determined utilizing free and not free lunches. The results were similar to those in earlier studies. Of 253 retained kindergarten students, 63.2% received free lunch, almost twice the number paying for lunch. Of the entire populuation, 48.4% received free lunch while 51.6% did not. Table 5 shows results. Significance again was evident.

Disaggregation of retained students entering STAR by socio-economic status for first graders produced results similar to those reported in earlier studies. Of the 1117 (of 1153) retainees who reported on free lunch, 69.2% were on free lunch and 30.8% were not on free lunch. Approximately one and a half times as many retained students received free lunch as those not receiving free lunch. The difference is significant. Of the entire population, 61% of the first graders received free lunch. See Table 5.

The distribution of retained kindergartners entering STAR by school type differed from that of many previous studies. Of the four school types, the largest percents of previously retained kindergarten students were in rural and suburban schools, with approximately 58% and 23% retained respectively as compared to 7% in inner-city and 12% in urban schools. There is a variation from the expected. Numbers of retained and not retained by school type appear in Table 6.

As with the kindergartners, the largest number of first grade retainees was found in rural schools and the least number in urban schools. Of the retained population, approximately 40% of the retentions occurred in rural schools. Only 9% were in urban compared to 24% in inner city and 26% in suburban schools. Of students entering STAR in grade one, more than half of those from rural areas (54.8%) and from inner-city schools (54.8%) had been retained in grade one (1985-86). For urban schools and suburban schools, the rate was slightly lower (52% and 42% respectively). Table 6 summarizes this information.

ACHIEVEMENT SCORES OF KINDERGARTEN AND FIRST GRADE GROUPS

KINDERGARTNERS AND THE STANFORD TEST

Table 7 summarizes mean scores of the reading section and math section of the Stanford test for those 253 students retained in kindergarten in 1984-85 and entering STAR in 1985-86 as second-time kindergartners. Of the 253 retained kindergarteners, 228 students (90%) took tests in K so that their progress was availbable to be followed from K through 3. This reduction in total number of students with test scores could be due to a second retention, a move out of STAR, placement in special education, or failure of students to take all tests. Continuation of the 228 retainees entering STAR in kindergarten who had K test scores using the largest number of test takers is as follows by class type: (S): 61 to 18 (30%); (R): 93 to 37 (40%); (RA): 77 to 20 (26%); total group: 228 to 75 (33%).

A comparative look at the scores in reading and math across four years shows that retainees in regular classes performed better than retainees in S and RA in all cases except one (retainees in S in math in K). Small class students did better than R and RA in only three cases--better than RA in reading by .8, better than R by 3.2 points in reading, and better than RA by 9.1 points in math, each in kindergarten. In all other cases, the test results of S class students fell behind those of RA students who sored lower than R class students. There is no significant difference between and within groups. The pattern of mean scores fails to reflect any remediation effect offered by the S condition for retained kindergarten students.

Of the original 6041 new kindergartners, test scores were available on 5617 (93%) in kindergarten. By the end of third grade, scores were available on 2845 of the students. Continuation of the largest group testing by class size breaks down as (S): 1694 to 898 (53%); (R): 1932 to 971 (50%); (RA): 1991 to 976 (49%);and (Total): 5617 to 2845 (51%).

A different pattern emerges when looking at the means of reading and math scores of non-retainees for four years. At every grade level in both reading and math, students in the small-class condition outscored those in R and RA by a significant margin. Additionally, these students outscored those second-time kindergartners in all three class sizes. In 67% of the cases, retained students showed less variance in their scores as compared to new kindergartners. This may be attributed to maturity in test taking. However, once retained, kindergartners were not able to catch up. Achievement scores for kindergartners not retained appear in Table 8. Effect sizes for the reading scores of the Stanford Achievement Test for kindergartners entering STAR in 1985 appear in Appendix B.

FIRST GRADERS AND THE STANFORD TEST

Of the 1152 retained first grade students, test data on 89% were available for their second time in grade one. By grade three, 53% of the 1152 students were tested. Continuation in STAR using the largest number of test takers by class size is as follows: (S): 153 to 66 (43%); (R): 238 to 505 (47%); (RA): 236 to 438 (54%); and total 1096 to 540 (49%).

As with the retained kindergarten students, generally no significant difference was found between and within groups for retained first graders. Only in grade one with math scores was there a significant difference between R and RA and again in grade two in reading between the same groups. The pattern of mean scores shows that no single class size made a difference to retained students.

Achievement scores for these students from first grade through third are summarized in Table 9.

The picture of achievement among students who entered STAR at age or who were not retained in grade one is not as clear as that of first time kindergartners. While students in S always outscored those in the other two conditions, the difference was only significant at grade one in reading and math and again in reading in grade two. There was also a significant difference between R and RA in reading and math in grade one and between R and RA in math in grade two. No statistical difference was found in grade three. This information is found in Table 10. Effect size for reading scores of the Stanford Achievement Test for first graders entering STAR in 1986 are in Appendix C.

Of the 1124 students identified as new first graders, test data were available in grae one for 1058 or 94%. Forty-five percent were tracked through grade three. Following students through STAR using the largest group looks like this: (S): 202 to 96 (48); (R): 519 to 188 (36%); RA: 408 to 107 (46%); and total 1076 to 471 (44%).

KINDERGARTNERS AND THE BASIC SKILLS FIRST TEST

Tables 11 and 12 show the percent passing the Basic Skills First Test (BSF) for those new entrants to kindergarten and those second-time kindergartners at the end of grades one, two, and three in reading and math. As with the SAT results, the kindergartners who had not been retained performed better in small classes than those in regular or regular with an aide in both reading and math. While the difference between groups is statistically significant, this is most likely due to the large number of students tested (3463). Additionally, no matter the class size, new kindergartners had a higher percentage passing than the retainees. Retention did not enable the retainees to catch up as so many educators propose.

Retained kindergartners in small class failed to perform as well as those in regular or regular with an aide. Table 11 shows that retainees had a lower percent passing in small class in both reading and math than did pupils in R and RA in each of the three grade levels. In grade one, retainees in RA have a higher percent passing in both reading and math than did pupils in R and S. This is true in grade two in math and in reading in grade three. Students in R have a higher percent passing in reading in grade two and in math in grade three than did pupils in either of the other two conditions. There is no statistical significance at $p \leq .05$. Again, once a child was retained, small class does not improve his scores.

FIRST GRADERS AND THE BASIC SKILLS FIRST TEST

The 1124 new entrants to first grade and the 1152 first grade retainees were tracked through third grade, this time on the Basic Skills First test. Percent passing for both groups in math and reading is given in Table 14 along with the numbers of those tested. In grade one on the math section of the test, the retainees had a higher percent passing the test only in the RA condition than did the new first graders. In all other cases, the new first graders outperformed the retained first graders.

As has been the pattern in this study, those students not retained performed better in small class. There did appear one exception to this at the third grade level in the math section of the test. Here the RA students had a higher percent passing, but there was no statistical significance at $p \leq .05$.

There was no statistical significance between groups for the retained first graders at any of the three grades. Yet, students in S did have a higher percent passing the test in reading and math in grades one and two, and in math in grade three. A difference of 2-4 points was found. This information is also in Table 13. Even with the slight variation in scores, there is no remedial effect evident from small class size with retained students.

The points gained on the SAT by kindergartners from K to 1, 1 to 2, and 2 to 3 show a pattern between retained and not retained groups. The new kindergartners had a larger point gain from K to 1. This gain decreased from 1 to 2, and by grade 3, the retainees showed a larger point gain than the group not retained. The first grade retainees did not surpass the group of new first graders. (See Tables 15 and 16.)

NUMBER OF STUDENTS TRACKED

Tables 15 and 16 give the number and percentage of students for both the retained and the non-retained kindergartners (1985 entrance into STAR) who were tracked from K-3 on the Stanford test and the Basic Skills First test. The largest group tested is traced. A larger percentage of students remained in S condition for these tests in the group of new entrants to kindergarten. Those in the S condition showed a larger percentage remaining in the study for retained kindergartners on the BSF but a larger group for the retained kindergartner in the R condition for the SAT.

Table 1

Placement by Class Size - Kindergarten

Class Type	Not Retained	Retained	Row Total
S			
n	1819	71	1890
Row %	96.2	3.8	100.0
Col %	30.1	28.1	30.0
R			
n	2085	99	2184
Row %	95.5	4.5	100.0
Col %	34.5	39.1	34.7
RA			
n	2137	83	2220
Row %	96.3	3.7	100.0
Col %	35.4	32.8	35.3
Column Total			
n	6041	253	6294
Row %	96.0	4.0	100.0
Col %	100.0	100.0	100.0

$X^2 = 2.28$

$p \leq .32$

Table 2

Placement by Class Size - Grade 1

Class Type	Not Retained	Retained	Row Total
S			
n	213	166	379
Row %	56.2	43.8	100.0
Col %	19.0	14.4	16.7
R			
n	489	522	1011
Row %	48.4	51.6	100.0
Col %	43.5	45.3	44.4
RA			
n	422	464	886
Row %	47.6	52.4	100.0
Col %	37.5	40.3	38.9
Column Total			
n	1124	1152	2276
Row %	49.4	50.6	100.0
Col %	100.0	100.0	100.0

$X^2=8.55$

$p \leq .01$

Table 3

Frequency of Sex of Kindergartners Retained and Not Retained Entering Star in 1985 and of First Graders Retained and Not Retained Entering Star in 1986

	Kindergarten[a]			1st Grade[b]		
	Retained	Not Retained	Row Total	Retained	Not Retained	Row Total
Male						
n	175	3060	3235	714	531	1245
Row%	5.4	94.6	100.0	57.3	42.7	100.0
Col%	69.2	50.7	51.4	62	47.2	54.7
Female						
n	78	2981	3059	438	593	1031
Row%	2.5	97.5	100.0	42.5	57.5	100.0
Col%	30.8	49.3	48.6	38.0	52.8	45.3
Column Total						
n	253	6041	6294	1152	1124	2276
Row%	4.0	96.0	100.0	49.4	50.6	100.0
Col%	100.0	100.0	100.0	100.0	100.0	100.0

[a] X^2=33.33; $p \leq 0.00$

[b] X^2=49.86; $p \leq 0.00$

Table 4

Frequency of Race of Kindergartners Retained and Not Retained Entering Star in 1985 and of First Graders Retained and Not Retained Entering Star in 1986

	Kindergarten[a]			1st Grade[b]		
	Retained	Not Retained	Row Total	Retained	Not Retained	Row Total
White						
n	201	4015	4216	702	661	1363
Row%	4.8	95.2	100.0	51.5	48.5	100.0
Col%	79.4	66.5	67	61	58.9	60
Non-White						
n	52	2026	2078	449	461	910
Row%	2.5	97.5	100.0	49.3	50.7	100.0
Col%	20.6	33.5	33.0	39.0	14.1	40.0
Column Total						
n	253	6041	6294	1151	1122	2273
Row%	4.0	96.0	100.0	50.6	49.4	100.0
Col%	100.0	100.0	100.0	100.0	100.0	100.0

[a] $X^2=18.51$; $p \leq 0.00$

[b] $X^2=1.02$; $p \leq 0.31$

Table 5

Frequency of Socioeconomic Status of Kindergartners Retained and Not Retained Entering Star in 1985 and of First Graders Retained and Not Retained Entering Star in 1986

	Kindergarten[a]			1st Grade[b]		
	Retained	Not Retained	Row Total	Retained	Not Retained	Row Total
Free Lunch						
n	160	2887	3047	773	574	1347
Row%	5.3	94.7	100.0	57.4	42.6	100.0
Col%	63.2	47.8	48.4	69.2	52.9	61.1
Not Free Lunch						
n	93	3154	3247	344	512	856
Row%	2.9	97.1	100.0	40.2	59.8	100.0
Col%	36.8	52.2	51.6	30.8	47.1	38.9
Column Total						
n	253	6041	6294	1117	1086	2203
Row%	4.0	96.0	100.0	50.7	49.3	100.0
Col%	100.0	100.0	100.0	100.0	100.0	100.0

[a] $X^2=23.21$; $p \leq 0.00$

[b] $X^2=61.95$; $p \leq 0.00$

Table 6

Frequency of School Type of Kindergartners Retained and Not Retained Entering Star in 1985 and of First Graders Retained and Not Retained Entering Star in 1986

	Kindergarten[a]			1st Grade[b]		
	Retained	Not Retained	Row Total	Retained	Not Retained	Row Total
Inner-City						
n	17	1403	3047	281	234	515
Row%	1.2	98.8	100.0	54.6	45.4	100.0
Col%	6.7	23.2	22.3	24.4	20.8	22.6
Suburban						
n	57	1347	1404	299	408	707
Row%	4.1	95.9	100.0	42.3	57.7	100.0
Col%	22.5	22.3	22.3	26.0	36.3	31.1
Rural						
n	148	2757	2905	465	383	848
Row%	5.1	94.9	100.0	54.8	45.2	100.0
Col%	58.5	45.6	46.2	40.4	34.1	37.3
Urban						
n	31	534	565	107	99	206
Row%	5.5	94.5	100.0	51.9	48.1	100.0
Col%	12.3	8.8	9.0	9.3	8.8	9.1
Column Total						
n	253	6041	6294	1152	1124	2276
Row%	4.0	96.0	100.0	50.6	49.4	100.0
Col%	100.0	100.0	100.0	100.0	100.0	100.0

[a] $X^2=41.18$; $p \leq 0.00$

[b] $X^2=28.99$; $p \leq 0.00$

Table 7

<u>Stanford Test Scores of Retained Kindergartners K-3</u>

	K				Grade 1				Grade 2				Grade 3			
Class	Read		Math		Read		Math		Read		Math		Read		Math	
Type	*n*	$\bar{X}$	*n*	$\bar{X}$	*n*	$\bar{X}$	*n*	$\bar{X}$	*n*	$\bar{X}$	*n*	$\bar{X}$	*n*	$\bar{X}$	*n*	$\bar{X}$
S	59	422.3	61	475.1	45	485.3	49	503.2	34	548.7	33	542.8	18	587.0	17	593.7
R	93	427.4	93	471.9	63	496.0	76	508.4	50	557.0	50	556.4	37	607.0	36	606.7
RA	76	421.5	77	466.0	41	486.8	47	503.4	35	551.8	36	546.3	20	604.1	20	602.9
Total	228		231		149		172		119		119		75		73	
F	1.89		1.04		0.74		0.33		0.34		1.20		1.52		0.79	
p	0.16		0.35		0.48		0.72		0.71		0.30		0.23		0.46	

Table 8

Stanford Test Scores of Kindergartners Not Retained K-3

	K				Grade 1				Grade 2				Grade 3			
Class	Read		Math		Read		Math		Read		Math		Read		Math	
Type	n	$\bar{X}$	n	$\bar{X}$	n	$\bar{X}$	n	$\bar{X}$	n	$\bar{X}$	n	$\bar{X}$	n	$\bar{X}$	n	$\bar{X}$
S	1673	441.2	1694	491.6	1292	536.0	1319	542.6	1027	598.8	1023	594.4	886	630.1	898	631.3
R	1906	435.1	1932	483.7	1393	525.3	1415	533.0	1112	594.1	1111	589.7	964	623.5	971	626.4
RA	1959	436.0	1991	483.4	1460	523.9	1502	532.2	1106	591.3	1104	586.8	960	622.6	976	625.6
Total	5538		5617		4145		4236		3245		3238		2810		2845	
F	18.97		16.64		18.91		24.59		7.51		7.92		11.33		5.50	
p	0.00		0.00		0.00		0.00		0.00		0.00		0.00		0.00	

Table 9

Stanford Test Scores of Retained First Graders (1-3)

Class	Grade 1				Grade 2				Grade 3			
	Read		Math		Read		Math		Read		Math	
Type	n	$\bar{X}$	n	$\bar{X}$	n	$\bar{X}$	n	$\bar{X}$	n	$\bar{X}$	n	$\bar{X}$
S	146	501.1	153	523.6	96	562.9	95	565.7	65	595.9	66	598.9
R	472	498.9	505	517.9	336	554.8	339	557.7	234	590.6	238	595.5
RA	405	506.5	438	523.3	297	561.2	296	561.3	228	596.2	236	598.7
Total	1023		1096		729		730		527		540	
F	2.70		2.67		2.67		1.64		1.75		0.50	
p	0.07		0.07		0.07		0.19		0.18		0.60	

Table 10

Stanford Test Scores of First Graders Not Retained (1-3)

Class Type	Grade 1 Read n	Grade 1 Read $\bar{X}$	Grade 1 Math n	Grade 1 Math $\bar{X}$	Grade 2 Read n	Grade 2 Read $\bar{X}$	Grade 2 Math n	Grade 2 Math $\bar{X}$	Grade 3 Read n	Grade 3 Read $\bar{X}$	Grade 3 Math n	Grade 3 Math $\bar{x}$
S	199	522.8	202	531.0	133	596.5	113	585.5	54	627.3	96	625.5
R	459	527.0	466	519.7	251	583.2	249	575.5	186	621.5	188	624.3
RA	400	517.5	408	525.5	243	596.7	242	585.2	184	622.7	187	624.9
Total	1058		1076		607		604		464		471	
F	8.12		5.64		5.19		4.20		0.77		0.03	
p	0.00		0.00		0.01		0.02		0.46		0.97	

Table 11

BSF Percent (Rounded) Passing By Grade (1-3) for Condition (S,R,RA) By Prior Retention in K, Star, 1989

	Grade 1				Grade 2				Grade 3			
	Read		Math		Read		Math		Read		Math	
Class Type	n	%	n	%	n	%	n	%	n	%	n	%
S	39	70	39	76	38	65	38	74	29	67	29	70
R	48	69	49	79	38	68	39	78	26	71	25	80
RA	44	73	45	83	43	67	44	81	32	74	33	75
Total	131	70	133	80	119	67	121	78	87	71	87	75
p	0.58		0.33		0.84		0.28		0.42		0.24	

Table 12

<u>BSF Percent (Rounded) Passing By Grade (1-3) for Condition (S,R,RA) By No Retention in K, Star, 1989</u>

Class	Grade 1				Grade 2				Grade 3			
	Read		Math		Read		Math		Read		Math	
Type	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
S	1208	88	1202	92	1128	87	1247	90	1103	85	1101	88
R	1161	84	1153	89	974	85	987	89	736	84	735	87
RA	1094	85	1091	90	1072	86	1093	90	987	84	986	86
Total	3463	86	3446	90	3274	86	3327	90	2826	84	2822	87
p	0.00		0.00*		0.00*		0.00*		0.05*		0.08*	

* Probably heavily influenced by the large *n*.

Table 13

BSF Percent (Rounded) Passing By Grade (1-3) for Condition (S,R,RA) Retained into First Grade, Star, 1989

Class Type	Grade 1				Grade 2				Grade 3			
	Read		Math		Read		Math		Read		Math	
	n	%	n	%	n	%	n	%	n	%	n	%
S	154	79	151	88	136	75	138	85	123	70	123	76
R	481	76	480	86	307	73	313	82	198	71	199	74
RA	438	78	435	86	314	72	324	81	255	71	258	74
Total	1073	77	1066	86	757	73	775	82	576	71	580	74
p	0.08		0.23		0.33		0.01		0.75		0.51	

Table 14

BSF Percent (Rounded) Passing By Grade (1-3) for Condition (S,R,RA) By No Retention in First Grade, Star, 1989

Class	Grade 1				Grade 2				Grade 3			
	Read		Math		Read		Math		Read		Math	
Type	n	%	n	%	n	%	n	%	n	%	n	%
S	194	85	190	91	144	86	145	90	138	84	137	86
R	455	80	454	86	213	82	212	87	136	83	136	85
RA	389	83	388	88	259	85	264	89	202	83	200	87
Total	1038	82	1032	88	616	84	621	87	476	83	473	86
p	0.00		0.00		0.01		0.03		0.51		0.70	

Table 15

Total Kindergartners Tested on Stanford Tests K-3

Class Type	Retained	Not Retained
S		
Number	61 to 18	1094 to 898
Percentage	30%	53%
A		
Number	93 to 37	1932 to 971
Percentage	40%	50%
RA		
Number	77 to 20	1991 to 976
Percentage	26%	49%

Table 16

Total Kindergartners Tested on Basic Skills K-3

Class Type	Retained	Not Retained
S		
Number	39 to 29	1208 to 1109
Percentage	74%	91%
A		
Number	49 to 26	1161 to 736
Percentage	53%	63%
RA		
Number	83 to 33	1094 to 987
Percentage	40%	90%

Table 17

Point Gains on the Stanford Test by Kindergartners

	K to 1		1 to 2		2 to 3	
Class Type	Retained	Not Retained	Retained	Not Retained	Retained	Not Retained
			Reading			
S	63	95	64	63	38	31
R	69	90	61	69	50	50
RA	65	88	65	67	52	32
			Math			
S	28	51	40	51	51	37
R	36	49	48	57	51	36
RA	37	49	43	55	57	39

Table 18

Point Gains on the Stanford Test by First Graders

	1 to 2		2 to 3	
Class Type	Retained	Not Retained	Retained	Not Retained
		Reading		
S	62	74	33	30
R	56	76	36	40
RA	54	76	35	29
		Math		
S	42	55	33	40
R	40	56	38	46
RA	38	59	38	40

CHAPTER V

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

CONCLUSIONS

This study was designed to address two questions: 1)what does the retained student look like in kindergarten and grade one, and 2)does class size make a difference in achievement of retained kindergartners and retained first graders. A portrait of the retained kindergarten is drawn from Project STAR as a white male from a low socio-economic background in a rural school. The population consisted of twice as many whites as minorities. The retained first grader was normally white, male, rural, with a low socio-economic background. Though an almost equal number of whites and non-whites were retained at this level, there were one and a half times as many whites as minorities in the population. Proportionately, the STAR retainee was a minority. Studies of this nature have normally portrayed the retainee as a male of minority race from inner-city schools. The STAR retainee matches the expected portrait of a retainee with the exception of school type. It is suggested that the geographic area of study might be a factor in retention.

In determining whether class size made a difference in achievement of retained kindergarten and first grade students, the findings from this study were conclusive. Tracking both retained kindergartners and retained first grade students through grade three, the emergent pattern showed that once a student had been retained, small class size failed to remediate test scores. Students who had not been retained consistently outscored those who had been held back regardless of class size. Findings from this study indicated that class size could not help a student once he or she had been retained.

QUESTIONS RAISED BY THE STUDY

Why did small class size not remediate test scores for those students who had been retained? Students in small class who had not been retained consistently outscored those in the other two conditions of R and RA. Futher research is suggested to answer this question. Additionally, why did the S condition generally show the largest percentage of students tracked K-3? Did the facilitative factors of the small class play a role here? Again, further research is needed. It is suggested that this study be expanded to continue tracking these two groups of students to determine longterm effects of class size.

IMPLICATIONS AND RECOMMENDATIONS

Schwager et al. (1992) summarized the status of retention:

> Retention has historically been seen as a solution to student failure. By controlling the flow of low-achieving students through a system of mass compulsory education, retention practices give the appearance of accountability and enforcement of standards without intervening in the underlying problem, that of low student achievement. As an organizational solution, retention is convenient: costs can be passed on to taxpayers through the general education budget and no change in system structure is required for implementation (p.435).

Educators in the United States must plead guilty as charged. Evidence has been overwhelming, both within this country and outside it. As has often been the case in education, America ignores the successful strategies of other countries. While the United States continues to propose retention as a means of strengthening standards and promoting stronger student performance, countries such as Denmark, Japan, Germany, Canada, and England do not employ retention as an instructional strategy in the elementary grades; their students out-perform ours (McAdams, 1993).

Politicians, policy makers, and educators alike stop to listen when money is involved. Perhaps it is from the monetary standpoint that these people must be approached, rather than one of test scores and self-esteem. It is evident from Project STAR and the Lasting Benefits Study that the cost of retention is higher in relation to small class size. A comparison of cost for retention and remediation in grade level was made earlier in this study. Surely, the triangulation of achievement, self esteem, and cost should serve to promote change in policy regarding retention and promotion.

Lao-tzu is credited with this bit of wisdom: "A journey of a thousand miles must begin with a single step." He states quite succinctly the philosophy educators must accept in light of the change many of their policies necessitate. Two such policies are those relating to retention and class size. The research is unquestionable in its conclusion that retention is not beneficial and it continues to accumulate on the side of small class size. Implications are many.

Practices of retention and large class size are not going to disappear over night. Change is arduous and best received in incremental steps. While retention policies exist, revisions must be made. Simultaneously, a re-educative program about retention and its effects and the benefits of small classes must occur. In conjunction, high quality programs and alternative strategies to retention must be investigated, developed, and implemented. These three prongs will form a comprehensive program designed to meet children's needs.

Policy on grade retention must be rethought. Retention decisions must consider many facets and be based on multiple criteria--criteria which is rational and standardized across districts (Cross, 1984;Sandoval,1984). Well-defined, specific criteria will promote equality in decisions made. Retention policy must be mutually developed by parents, teachers, and administrators. Staff development should be provided to ensure an understanding of the policy with staff interpretations examined. According to Schwager et al. (1992), retention policies are designed to respond to organizational symptoms of low classroom performance; it is important to look at how teachers respond to the underlying problems of low achievement, academic competency, and problematic behavior rather than simply assessing their impact on the frequency with which students are retained. It is equally as important to have constant input from students, teachers, and parents on the effects of the policy.

While policy is being reworked, a strong re-educative program about the effects of retention and small class size should be underway. It is the responsibility of educational professionals to be adequately informed on what the retention and class size data indicate. Both formal and informal presentations should be given. Parents should be invited and encouraged to attend similar sessions.

Districts should then become familiar with their own practices and provide an analysis at additional presentations. Information should be shared with legislatures and school boards. It is imperative that the policy makers understand fully the impact of their regulations.

Neither of these recommendations will be easily implemented. It is within the third domain that the most progress has already been made, albeit insufficient. All alternatives to assist low achieving students should be explored as should alternative strategies to retention and large classes. The literature is brimming with suggestions to escape retention, but is less than replete with recommendations to provide teachers with the smaller caseloads that would help alleviate the symptoms of retention.

Districts across the United States pilot numerous programs to improve instruction and learning yearly. Outcome-based education and nongraded elementary schools are currently on the forefront but date back to the origins of school. Ostrowski (1987) viewed this strategy as alleviating the need for retention\promotion policies based on values identified by Goodlad and Anderson (1964):

> The nongraded school provides for the continuous, unbroken, upward progression of all pupils, the slowest and the most able. The nongraded school provides for the irregular upward progression that is characteristic of almost every child. The nongraded school provides several alternative vertical classroom placements for every child at any time, no one of which denotes nonpromotion or skipping (p.35).

Districts should continue their search for worthwhile designs. Summer school programs have been joined by morning, afternoon, and weekend tutorials. Peer tutoring accompanies adult tutoring. Pre-school and full day kindergarten programs are offered for at-risk youngsters. Research continues on half step and transitional plans. Programs to strengthen parent involvement can be instrumental as will programs that better educate parents in child development. Implementation of flexible standards of competency and delay of achievement testing until after grade three are also strong strategies.

Additional recommendations concentrate on remediation within grade and teacher preparation. The causes of student failure must be assessed and those causes addressed. Bloom (1981) stated: "Failure of children to succeed with learning tasks should be regarded as a failure of curriculum and instruction rather than as a failure of the children" (108). Training early childhood teachers to develop and use child-centered, developmentally appropriate programs that meet the students' learning needs is an essential step in the program to eliminate retention. Teachers must also be properly trained to diagnose individual difficulties and provide instruction accordingly. Techniques such as cooperative learning, hands-on learning, learning styles, and performance assessment should be part of every teacher's training. A concentration of such strategies at the elementary level will soon become preventive measures of retention.

The need for small classes must also be addressed as a preventive measure against retention. Research on class size should be reported along side research on retention to educators, parents, and legislatures. The researchers from the STAR project view class size research not as an effort to reduce class size but as an attempt to find appropriate casework loads that permit the individual instruction and tutoring required by students (Pate-Bain, Achilles, Boyd-Zaharias, & McKenna, 1992). Bloom identified the most effective instruction as one-on-one tutoring; class size studies attempt to find a prudent alternative (Pate-Bain, Achilles, Boyd-Zaharias, & McKenna, 1992).

Further recommendations include additional research in the areas of class size and retention. Students from the STAR Project should continue to be tracked to determine if the gap between retained students and those students not retained continues to narrow. An analysis of retention policies across the nation could be conducted with the intent of identifying similar elements in the decision-making process. A list of standards for creating policy could be developed as well. Continued study in the area of class size as it relates to student achievement, teaching styles, learning styles, and retention is suggested. High retention rates are symptomatic of the real illness of low classroom performance. Are there facilitative factors which act to reduce retention and increase student learning and achievement? What are those factors? How and why does class size affect retention, teaching, and learning? This information cannot be gained through a post hoc study. Ethnographic methodology linked with quantitative methodology could add much to the current body of research in this area.

Regardless of the recommendation, educators must accept that the responsibility of failure is their own; the blame can no longer be placed on the child. Policy makers must recognize that panaceas in education do not exist and that any ingredient in the remedy for the ills is expensive. It is no longer a question of whether additional costs can be incurred but at what point funds should be provided. Students cannot continue to be discriminated against in the name of higher standards. Educators have a duty to meet the needs of every child. Retention and large classes prevent the fulfillment of this duty.

BIBLIOGRAPHY

Abidin, R.R., Golladay, W.M., & Howerton, A.L. (1971). Elementary school retention: An unjustifiable, discriminatory and noxious educational policy. Journal of School Psychology, 9, 410-417.

Achilles, C.M., & Moore, C. (1986). Some comments on class size: A review of major issues. Draft.

Baenen, N.R. (1988). Perspectives after five years--Has grade retention passed or failed? Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 5-9, 1988). ED 300 424.

Billman, J. (1988). Two years of kindergarten: Ethical and curricular considerations. Position paper. ED 297 884.

Bossing, L., & Brien, P. (1980). A review of the elementary school promotion/retention dilemma. ED 212 362.

Bredekamp, S., & Shepard, L.A. (1989). How best to protect children from inappropriate school expectations, practices, and policies. In R. Reitz (Ed.), Retention in grade: Looking for alternatives, (pp.221-232). Bloomington: IN: Phi Delta Kappan Educationa Association.

Byrnes, D.A., & Yamamoto, K. (1985). Academic retention of elementary pupils: An inside look. Education, 106, 208-214.

Byrnes, D.A., & Yamamoto, K. (1986). Views on grade repetition. Journal of Research and Development in Education, 20, 14-20.

Carter, E.F. (1990). Decreasing the incidence of first grade retention through a diagnostic-prescriptive intervention program and teacher in-service. Practicum paper, Nova University. ED 327 341.

Chansky, N.M. (1964). Progress of promoted and repeating grade 1 failures. Journal of Experimental Education, 32, 225-237.

Cooley, William, & Bickel, William. (1986). Decision-oriented educational research. Boston: Kluwer-Nijhoff, 1-297.

Cross, Ray. (1984). Techer decision-making in student retention. Paper presented at the annual meeting of the American educational research association (New Orleans, LA, April 23-27, 1984). ED 252 930.

Dobbs, b., & Neville, D. (1967). The effect of nonpromotion on the achievement of groups matched from retained first graders and premoted second graders. The Journal of Educational Research, 60, 427-474.

Doyle, R. (1989). The resistance of conventional wisdom to research evidence: The case of retention in grade. Phi Delta Kappan, 71, 215-220.

Ellwein, M.C., Walsh, D.J., Eads II, G.M., & Miller, A. (1991). Using readiness tests to route kindergarten students: The snarled intersection of psychometrics, policy, and practice. Educational Evaluation and Policy Analysis, 13, 159-175.

Finlayson, H. (1975). The effect of nonpromotion upon the self-concept of pupils in primary grades. ED 155-566.

Finn, J.D., & Achilles, C.M. (1990). Answers and questions about class size: A statewide experiment. American Educational Research Journal, 27, 557-577.

Folger, J. (1989). Project STAR and class size policy. In J. Folger (Ed.), Peabody Journal of Education, 67, (pp.1-33).

Frymier, J. (1989). A study of students at risk: Collaborating to do research. Bloomington, IN: Phi Delta Kappa Education Association.

Frymier, J. (1990). A tale of two crises. Principal, 69, 52-53.

Gallup, G.H. (1986). The 18th annual Gallup poll of the public's attitudes toward the public the public schools. Phi DeltaKappan, 65, 33-59.

Godfrey, E. (1972). The tragedy of failure. Education Digest, 37, 34-35.

Goodlad, J. (1954). Some effects of promotion and nonpromotion upon the social and personal adjustment of children. Journal of Experimental Education, 22, 301-327.

Goodlad, J., & Anderson, R.H. (1963). The nongraded elementary school. New York: Harcourt, Brace & World, Inc. 1-60.

Grissom, J.B., & Shepard, L.A. (1989). Repeating and dropping out of school. In L.A. Shepard & M.L. Smith (Eds.), Flunking grades: Research and policies on retention. (pp.34-63). London: Falmer.

Hedges,L.V., & Stock, W. (1983). The effects of class size: An examination of rival hypotheses. American Educational Research Journal, 20, 63-86.

Holmes, C.T. (1989). Grade level retention effects: A meta analysis of research studies. In L.A. Shepard & M.L. Smith (Eds.), Flunking grades: Research and policies on retention. (pp.6-33). London: Falmer.

Holmes, C.T. (1983). The fourth r: Retention. Journal of Research and Development in Education, 71, 1-6.

Holmes, C.T., & Matthews, K.M. (1984). The effects of nonpromotion on elementary and junior high school pupils: A meta analysis. Review of Educational Research, 54, 225- 236.

House, E.R. (1989). Policy implications of retention research. In L.A. Shepard & M.L. Smith (Eds.), Flunking grades: Research and policies on retention, (pp.202-213). London: Falmer.

Jackson, G.B. (1975). The research evidence on the effects of grade retention. Review of Educational Research, 45, 613-635.

Jawarski, A.P., & Riley, S.M. (1985). When curriculum does not match readiness. Paper presented at the annual meeting of the American Educational Research Association (69th, Chicago, IL, March 31-April 4, 1985), ED 265 947.

Johnson, E.R., et al. (1990). The effects of early grade retention on the academic achievement of fourth-grade students. Psychology in the Schools, 27, 333-38.

Karweit, N.L. (1991). Repeating a grade: Time to grow or denial of opportunity? Report No. 16, Center for Research on Effective Schooling for Disadvantaged Students, Baltimore, MD. ED 336 493.

Mantzicopoulos, P., & Morrison, D. (1992). Kindergarten retention: Academic and behavioral outcomes through the end of second grade. American Educational Research Journal, 29, 182-198.

McDaniel, A. (1986). The long-term effects of pupil progression policy on academic achievement. Paper presented at the Annual Meeting of the American Educational Research Association (67th, San Francisco, CA, April 16-20, 1986). ED 278 088.

Mitchell, D., Carson, C., Badarak, G. (1989). How changing class size affects classrooms and students. Riverside: University of California, California Educational Research Cooperative.

Morrison, I.E., & Perry, I.F. (1956). Acceptance of overage children by their classmates. Elementary School Journal, 56, 217-220.

Nason, Beth. (1991). Retaining children. Is it the right decision? Childhood Education, 300-304.

Newman, J.A. (1988). Retention/promotion/transition in the early grades: A research brief. Educational Service District 189, Mt. Vernon, Wash., 1-9. ED 303 278.

Nikalson, L.B. (1987). Do certain groups of children profit from a grade retention? Psychology in the Schools, 24, 339-345.

Nikalson, L.B. (1984). Nonpromotion: A pseudoscientific solution. Psychology in the Schools, 21, 485-499.

Norton, M.S. (1990). Practical alternatives to student retention. Contemporary Education, 61, 204-208.

Nye, B.A., Achilles, C.M., Zaharias, J.B., Fulton, B.D., & Wallenhorst, M.P. (1992). Small is far better. Paper submitted for MSERA Outstanding Paper Competition in Knoxville, Tennessee, November, 1992).

Ostrowski, P.M. (1987). Twice in one grade = a false solution. A review of the pedagogical practice of grade retention in elementary schools: What do we know? Should the practice continue? Report for Rhode Island State Dept. of Education, Providence. ED 300 119.

Owens, S.A., & Ranick, D.L. (1977). The Greensville program: A commonsense approach to basics. Phi Delta Kappan, 58, 531- 533.

Pate-Bain, H., Achilles, C.M., Boyd-Zaharias, J., & McKenna, B. (1992). Phi Delta Kappan, 74, 253-256.

Peck, P. (1989). In search of solutions. Instructor, 98, 28-30.

Peterson, S.E., DeGracie, J.S., & Ayabe, C.R. (1987). A longitudinal study of the effects of retention/promotion on academic achievement. American Educational Research Journal, 24, 107-118.

Plummer, S.L., et al. (1984). The academic and social consequences of grade retention: A convergent analysis. ERIC Clearinghouse on Elementary and Early Childhood Education, Urbana, Ill. ED 247 033.

Putka, G. (November 30, 1988). Getting tough: Schools start flunking many more students in drive on mediocrity. The Wall Street Journal, LXXXII, 1.

Reynolds, A.J. (1992). Grade retention and school adjustment: An explanatory analysis. Educational Evaluation and Policy Analysis, 14, 101-121.

Rose, J.S., Medway, F.J., Cantrell, V.L., Marus, S.H. (1983). A fresh look at the retention-promotion controversy. Journal of School Psychology, 21, 457-462.

Routh, Nancy. (1986). Effect of grade retention on student achievement: a case study of the effectiveness of grade retention in improving the achievement of elementary students. Dissertation, University of North Carolina at Greensboro.

Sandoval, J. (1984). Repeating first grade: How the decision is made. Psychology in the Schools, 21, 457-462.

Schultz, T. (1989). Testing and retention of young children: Moving from controversy to reform. Phi Delta Kappan, 71, 125-129.

Schuyler, N.B. Does retention help? Perspectives after three years. Paper presented at the Annual Meeting of the American Educational Research Association (69th, Chicago, IL, March 31-April 4, 1985), ED 263 143.

Schwager, M.T., Mitchell, D.E., Mitchell, T.K., & Hecht, J.B. (1992). How school district policy influences grade level retention in elementary schools. Educational Evaluation and Policy Analysis, 14, 421-438.

Shepard, L.A. (1989). A review of research on kindergarten retention. In L.A. Shepard & M.L. Smith (Eds.), In Flunking grades: Research and policies on retention. (pp.202-213). London: Falmer.

Shepard, L.A., & Smith, M.L. (1987). Effects of kindergarten retention at the end of first grade. Psychology in the Schools, 24, 346-357.

Shepard, L.A., & Smith, M.L. (1988). Escalating academic demand in kindergarten: Counterproductive policies. The Elementary School Journal, 89, 134-145.

Shepard, L.A., & Smith, M.L. (1989). Flunking grades: Research and policies on retention. London: Falmer.

Shepard. L.A., & Smith M.L. (1986). Synthesis of research on school readiness and kindergarten retention. Educational Leadership, 44, 78-86.

Slavin, R.E. (in press). Achievement effects of substantial reductions in class size. School and Classroom Organization. Hillsdale, NJ: Erlbaum.

Slavin, R.E., & Karweit, N.L. (1992). Preventing early school failure: What works? Educational Leadership, 50, 10-18.

Smith, M.L., & Shepard, L.A. (1988). Kindergarten readiness and retention: A qualitative study of teachers' beliefs and practices. American Educational Research Journal, 25, 307-333.

Smith, M.L., & Shepard, L.A. (1987). What doesn't work: Explaining policies of retention in the early grades. Phi Delta Kappan, 69, 129-134.

Stroup, Stinson W., & Zirkel, Perry A. (1983). A legal look at the retention/promotion controversy. Journal of school psychology, 21, 213-217.

Tomchin, E.M., & Impara, J.C. (1992). Unraveling teachers' beliefs about grade retention. American Educational Research Journal, 29, 199-223.

Walden, J.C., & Gamble, L.R. (1985). Student promotion and retention policies: Legal considerations. In R. Reitz (Ed.), Retention in grade: Looking for alternatives, (pp.201-216). Bloomington, IN: Phi Delta Kappan Educational Association.

Word, E., Johnston, J., Bain, H., Fulton, B., Zaharias, J., Lintz, N., Achilles, C.M., Folger, J., & Breda, C. (1990). Student/teacher achievement ration (STAR): Tennessee's K-3 class size study. Final report and final report summary. Nashville, TN: Tennessee State Department of Education.

APPENDIX A

STAR Design and Results
Primary and Extended Analyses Designs: STAR (1985-1989)

Sample Design:

4 Locations (Urban, rural, etc.)	(Fixed Effed)
Schools nested in Locations	(Random Effect)
Class types (S,R, RA) crossed with locations and school types	(Fixed Effect)
2 Training categories	(Fixed)

Source Table

Source of Variation:	Error Term:
Location (L)	Schools
Training (TR)	Schools
Type T	School x type
LxT	School x type
LxTR	School
TxTR	School x type
LxTxTR	School x type

		Degrees of freedom (df) Ach. Meas.	Noncog. Meas.
Schools	e.g. (1986)	75	69
School x Type	e.g. (1987)	149	137
Classes within School-Types (etc.)		Etc.	

Primary Model: Measures		Matched
Achievement (Ach):	SESAT, SAT, BSF	t-tests
Noncognitive (Noncog):	SCAMIN, Attendance, behavior, etc.	

Extended Model: Measures:
Sex (or Race, or SES) Ave. Diff Scores on Ach. Multivariate
Ses (or Race, or SES) Ave. Diff Scores on Noncog. Models
Training

Two planned contrasts: S class mean vs means of all R and RA; S vs (r + Ra 2); RA classs mean vs R class mean.

Analysis of Variance for Cognitive Outcomes, STAR, Grades K-3. Sig. Levels p<.05 or Greater are Tabled.

Effect/ Grade		Reading Multi- variate	Reading SAT Read	Reading BSF Read	Mathematics Multi- variate	Mathematics SAT Math	Mathematics BSF Math
Location (L)	K		.02			.05	
	1	.01	.06		.05		
	2	.001	.001	.001		.001	.001
	3	.001	.001	.001	.001	.001	.001
Race (R)	1	.001	.001	.001	.001	.001	.001
	2	.001	.001	.001	.001	.001	.001
Type (T)	K		.001			.02	
	1	.001	.001	.001	.001	.001	.05
	2	.001	.001	.05	.001	.001	.05
	3	.001	.001	.001	.001	.001	.001
SES	K		.001			.02	
Loc X Race	1	.05		.05			
Loc X Type	K-3	All N/S. The class-size effect is found in all locations--Inner City, Suburban, Urban, and Rural schools. (Tabled as important).					
Race X Type	1	.05	.05	.01			
LxRxT	1			.05			.01
LxTRxT	2	.05	.01	.05	.05	.05	.01

APPENDIX B

Effect Size For the Reading Section of SAT Kindergartners Entering Star in 1985

	Retained (n=253)			Not Retained (n=6041)				
Class Type	n	SD	$\bar{X}$	n	SD	$\bar{X}$	Difference	Effect Size
				K				
S	59	19.53	422.25	1673	32.73	441.21	18.96	.60
R	93	27.17	427.39	1906	31.10	435.06	7.67	.25
RA	76	15.94	421.25	1959	31.86	435.98	14.73	.46
Total	228	22.10	424.01	5538	31.97	437.24	13.23	.41
				1st				
S	45	45.26	485.31	1292	57.30	536.03	50.72	.89
R	63	59.83	496.05	1393	54.97	525.31	29.26	.53
RA	41	35.71	486.80	1460	55.59	523.87	37.07	.67
Total	149	49.78	490.26	4145	56.17	528.15	37.89	.67
				2nd				
S	34	48.09	548.74	1027	44.73	598.80	50.06	1.12
R	50	48.88	557.02	1112	44.89	594.14	37.12	.83
RA	35	41.46	551.77	1106	44.50	591.35	39.58	.89
Total	119	46.34	553.11	3245	44.80	594.66	41.55	.93
				3rd				
S	18	46.79	587.00	886	37.23	630.09	43.09	1.16
R	37	42.13	607.03	964	36.43	623.46	16.43	.45
RA	20	31.38	604.05	960	37.02	622.56	18.51	.50
Total	75	41.08	601.43	2810	37.02	625.24	23.81	.64

APPENDIX C

Effect Size For the Reading Section of SAT First Graders Entering Star in 1986

Class Type	Retained (n=1153) n	SD	$\bar{X}$	Not Retained (n=1124) n	SD	$\bar{X}$	Difference	Effect Size
				1st				
S	146	45.28	501.12	199	52.77	522.76	21.64	.41
R	472	49.26	498.95	459	51.02	507.00	8.05	.16
RA	405	48.28	506.48	400	50.88	517.53	11.05	.22
Total	1023	48.40	502.24	1058	51.65	513.94	11.70	.23
				2nd				
S	96	41.24	562.94	113	44.71	596.54	33.6	.75
R	336	38.78	554.78	251	45.58	583.20	28.42	.62
RA	297	42.10	561.24	243	41.04	593.72	32.48	.79
Total	729	40.58	558.49	607	43.95	589.90	31.41	.71
				3rd				
S	65	34.13	595.86	94	34.65	627.33	31.47	.91
R	234	33.24	590.65	186	40.00	621.52	30.87	.77
RA	228	33.56	596.21	184	36.77	622.65	26.44	.72
Total	527	33.54	593.70	464	37.60	623.15	29.45	.78