INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University
Microfilms
International

300 N. Zeeb Road
Ann Arbor, MI 48106

Maddocks, John Calvin

THE RELATIONSHIP BETWEEN AVERAGE STUDENT ACHIEVEMENT AND NONPROMOTION RATE: A PATH ANALYSIS MODEL FOR NORTH CAROLINA ELEMENTARY SCHOOLS

The University of North Carolina at Greensboro ED.D. 1983

THE RELATIONSHIP BETWEEN AVERAGE STUDENT ACHIEVEMENT AND NONPROMOTION RATE: A PATH ANALYSIS MODEL FOR NORTH CAROLINA ELEMENTARY SCHOOLS

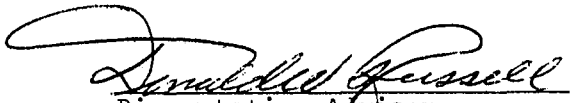
by

John C. Maddocks

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
1983

Approved by



Dissertation Advisor

APPROVAL PAGE

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

Dissertation Advisor _____

Committee Members _____

_____

_____

December 20, 1982
Date of Acceptance by Committee

December 20, 1982
Date of Final Oral Examination

MADDOCKS, JOHN CALVIN. The Relationship Between Average Student Achievement and Nonpromotion Rates: A Path-Analysis Model for North Carolina Elementary Schools. (1983) Directed by: Dr. Donald W. Russell. Pp. 94.

This research examines the relationship among initial elementary school achievement, nonpromotion, and subsequent academic achievement. School-wide achievement means and retention rates were examined for first-, second-, and third-grade classes across North Carolina. In addition, the factors of race, socioeconomic status, and intelligence were examined as correlates of achievement means and nonpromotion rates and their linear effects were controlled in the analysis. This research analyzes the effects of differing rates of nonpromotion upon classes of children in several grades rather than the effects upon individual children who are either promoted or retained.

A review of previous research was conducted and these findings were compiled into an hypothesized path-analysis model. This model proposed a relationship among several variables, indicated causal links, and arranged them in temporal order.

Data were collected from three agencies of the North Carolina Department of Public Instruction and from nine local school administrative units across the state. Regression analysis was performed and the resulting standardized coefficients entered in the model as path coefficients. A replication of the reading achievement model was performed using math achievement data. Both models were reduced to a parsimonious form and these forms were examined for reasonableness. The stability of the data was examined for the three years under investigation.

The results support the contention that, on average, schools which retain a higher percentage of pupils in the first grade demonstrate higher mean achievement in the second grade. However, the magnitude of the effect is relatively small. First-grade achievement has almost three times the direct effect upon second-grade achievement as that of first-grade nonpromotion. The effect of nonpromotion on subsequent achievement was not confirmed beyond the second-grade level. Race and socioeconomic status had no effect upon school mean achievement or on non-promotion rates, except through their significant correlation with intelligence test scores. Separate models for reading and math achievement were found to be reasonably similar. The data were found to be stable during the period of investigation, and representative of the population of North Carolina public schools.

Unexamined school-wide and teacher variables appear to affect the rate of nonpromotion. Future replications of this study should examine these factors, as well as expand the scope of investigation to include intermediate and secondary grades.

ACKNOWLEDGEMENTS

I wish to thank Dr. Richard Jaeger, Dr. Dale Brubaker, Dr. Ernest Lumsden, and especially Dr. Donald Russell, chairman, for their willingness to serve on my doctoral committee and for their guidance along the way.

I wish to thank Dr. Robert Evans and the State Department of Public Instruction for their assistance in data collection.

I wish to thank my wife, Jane, and my children, Jenny, Jay, and Elizabeth, for their patience and support during this program of study.

TABLE OF CONTENTS

Page

CHAPTER

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION AND BACKGROUND

History of Grade Retention

During the early nineteenth century, the one-room school predominated throughout the United States. Small, scattered schools proliferated, necessitated by limited means of transporting pupils. Most children walked to school, thus limiting the geographic area any one school could cover. Since much land was used by small family-farm operations, population density remained low. These two factors resulted in a preponderance of single-room school houses.

As rural areas grew and small urban centers developed, the pupil population increased. Improvements in roads and transportation brought students from more remote settings into existing schools. These changes resulted in the consolidation of some small schools into larger, multi-room schools and the addition of second or third rooms to other single-room facilities. However, even where schools of two or three classes developed, their programs continued to be taught to students of varied ages. Some had each teacher instructing all ages; others divided teaching responsibilities by subjects; still others divided pupils by broad categories of competency such as readers and nonreaders. Grouping by age was rarely considered, for few schools had enough children of one age to form a class. Academic progress remained an individual matter because there were not enough children of a similar age assembled in one school to establish group expectations.

However, this did not mean there was an absence of failure. Children might fail a subject or fail to learn, but they were not retained in a grade and did not change groups of children. They continued to progress at their own rate within the same group. Failure to learn and failure to pass were always present, but failure to pass a grade and the concept of grade retention came into being with the development of the graded school in the mid-nineteenth century (Clark, 1963). Thus, significant differences in the interpretation of "failure" began to develop (Ebel, 1980).

The graded school first appeared in cities and later in rural areas. Multi-roomed schools were constructed wherein children were groups by age. Children could then be evaluated according to an age norm. Those who failed to make adequate progress as measured against their group's norm were retained for a repeat of the entire grade, not just a review of the material they had failed. Grade retention increasingly became the practice from the mid-1800's to the early 1900's at which time one in six pupils was retained each year. However, from 1909 to 1949, the national rate of retentions dropped from approximately 16 percent to around 5 percent (Clark, 1963). This decrease was the result of overwhelming criticism in the literature of the practice of retaining children in a grade. There appeared to be two main concerns. The first was a sincere interest in children's social and emotional well-being, and a desire to protect them from feelings of incompetence and insecurity. Alongside this was concern for the financial burden placed upon society by the added year and added costs of retaining pupils.

The research from 1910 until 1940 was uniformly against nonpromotions, citing pupil depression, discouragement, distrust of himself and his own

ability, and behavioral problems such as increased cruelty and bullying of other children (Scott & Ames, 1969). Cost and social factors were also noted. It was certainly a burden on the taxpayers to have a high percentage of failures. Moreover, the achievement of the retained pupils became a concern. Research during this period noted that most pupils who had been retained showed little improvement and frequently demonstrated lower achievement the second time around. Considering all these factors, it was not unusual that the nation's rate of retention should drop during the first half of the 20th century.

North Carolina has followed the trend demonstrated throughout the rest of the country. During the 1930's North Carolina retained one of every five students each year. However, by the 1960's the ratio of retained pupils had dropped to one in twenty (Clark, 1963). At the same time, the variability of ages within each grade also decreased. During the '50's and '60's, the promotion-retention rates remained relatively constant or dropped only slightly. North Carolina's rate of retention reached an all-time low during the early '70's. Since that time, it has been increasing slightly (NCSDPI, 1975; 1980).

The more recent swing toward increased retentions can be seen in more frequent literary references to the benefits of grade repetition. Scattered studies during the last 30 years have offered fewer negative viewpoints concerning the practice of retaining pupils. Less concern about the social and emotional impact upon pupils has emerged, and some studies even indicated increased academic performance following retention. Increased public concern about declining SAT scores during the late 60's and early 70's added pressure on educators to retain pupils who

demonstrated little academic progress. An increasing number of studies on testing to control grade-to-grade promotion were seen in the literature (Haney & Madaus, 1978). The competency testing movement was viewed as renewed emphasis on improved academic achievement. Curriculum reforms went hand in hand with increased pressure to retain nonachieving students. Beginning in 1950, studies began to indicate that careful selection of children repeating a grade could bring about much higher success rates than previously available data would have suggested (Lobdell, 1954). School policies established definite criteria for making promotion-retention decisions. One study indicated that social and personal adjustment of low-achieving repeaters appeared as good, if not better, than promoted low-achievers (Worth, 1960). Occasionally, retention was based upon sound policy and carefully selected students. However, caution was still indicated, for grade repetition was not a cure-all. It would not make successful those pupils of low intelligence, or those who were emotionally disturbed, or brain-damaged. It was felt retention might prove productive for an immature child, i.e., a child behaviorally or developmentally below his age (Scott & Ames, 1969). However, earlier research had lumped all retainees together.

Need for the Study

This study is both timely and necessary. The effects of student retentions have been argued for 75 years with little resolution. The vase majority of statistical research has been ambiguous and inconclusive. Some authors have found benefits in retaining pupils; others reported significant negative results. Still others reviewing

educational research conclude that many of the "findings" have been based on biased statistical procedures and faulty designs.

Most of the research has examined the effects of failure upon the child who was retained, or who might have been retained and was not. Little consideration has been given to the effects of retention upon the class, school, or school system. While one must never lose sight of the individual, concern for the educational process is also appropriate. It is not unreasonable to assume that failure affects the retained child differently than the nonretained child. Since the vast majority of children are not retained in any given grade, the net result of grade assignment policies may be significantly different than the effect upon the retained child.

What should be the policy with regard to children's repeating grades? Arguments for and against grade assignment plans are based not only upon sound research, but also emotionality, educational mythology, management concerns, and financial realities. While all promotion-retention theories profess concern for the child or for the system, contradictory policies have been founded upon contradictory theories. Thus, retention policy remains an unresolved issue affecting countless children at a major annual cost to the public. It is a contemporary problem in that the pendulum of educational practice is now swinging toward the conservative. The back-to-basics movement, management by objectives, and strict retention policies have common roots in conservative educational philosophies. It is now time for educators to examine the practice of failing students from a new perspective. This research examines retention and its relationship to school-wide achievement rates. Improved academic achievement is an educational goal espoused by theorists on both sides of this issue.

A demonstrated positive relationship between achievement means and retention rate would be consistent with the argument that risk of failure motivates pupil achievement. A significant negative relationship between the two would be consistent with the argument that failure damages the child, and thus, the class, which is composed of individuals.

There is a demonstrated need for more information, based on sound theoretical precepts, supported by careful analysis of valid data. This research provides information necessary for establishing or updating grade assignment policies.

Statement of the Problem

Although the practice of grade retention has been researched and discussed extensively for three quarters of a century, the effect of retention policies on school-wide performance has not been examined statistically. Suppositional writings have both lauded and condemned the possible effects of policies of failure upon the climate of our nation's schools. Some writers have asserted that failure is destructive of a pupil's self-image, resulting in adverse social, emotional, and behavioral adjustments. Thus it follows that schools supporting strict policies of promotion and retention adversely affect their school-wide social, emotional, and behavioral climate. Such an unhealthy environment might also affect the academic achievement of pupils who are not retained. Conversely, other authors have argued that establishing strict policies of grade assignment is indicative of a school's performance standard. While agreeing that some students may suffer as a result of failing, the proponents of strict policies contend that overall pupil achievement will rise because the

structure of predetermined standards establishes performance objectives and because achievement is rewarded while nonachievement is punished.

Both of these positions share three basic assumptions: they assume that retention rate is related to grade assignment policies and that higher rates of retention will exist where a stricter policy is in force. They assume that retention rates do effect the academic environment of the school. Finally, they accept as fundamental that the academic environment has an effect upon pupil achievement and, therefore, achievement rates. Therefore, it may be hypothesized that there exists a relationship between pupil retention rates and school-wide achievement rates. If a significant positive relationship were demonstrated between a school's mean score on a valid, standardized achievement test and its percentage of pupils retained in grade, it would be consistent with the contention that strict policies of pupil assignment improve mean academic achievement. If a significant negative relationship were demonstrated between these factors, it would be consistent with those theories arguing that retention hurts individuals, and, therefore, damages the mean academic achievement of all pupils.

Both theoretical positions assume that if a relationship is demonstrated between these two factors, that achievement is the dependent variable; that is, changes in academic performance are, in part, the result of alterations in retention rates. However, it might be argued as convincingly that changes in the rate of retention result from alterations in the mean achievement scores of a school. If pupil achievement were to be improved, for whatever reason, certainly fewer pupils would be retained.

This study attempts an analysis of the relationship between a school's retention rate and its mean pupil achievement, and examines the ambiguity in the direction of possible causality between these variables. No attempt is made to analyze the progress or lack of progress of individual children as a result of their retention or promotion, but rather, this research focuses on average academic achievement as it is related to a school's percentage of nonpromoted students. By examination of data in a longitudinal fashion, the effect of previous achievement scores upon retention can be examined, in addition to the effect of retention upon subsequent achievement. The following statistical model offers a theoretical framework for this "chicken-or-egg" controversy as well as considering the relationship among these and other variables assumed to affect promotion and achievement.

Summary of Procedures

The essential hypotheses suggest simple regression analysis. An analysis of mean grade achievement and percentage of pupils not promoted could be made by using the linear equation

$$Y = B_o + B_1X_1 + \epsilon$$

where: Y is mean achievement per school for a given grade,

X_1 is the percentage of nonpromotions per school for a given grade,

B_o is a constant related to achievement and not dependent on promotion,

B_1 is a factor representing both the direction and magnitude of effect of nonpromotion on achievement,

ϵ is the unexplained random component of Y.

Whereas this formula is mathematically sufficient for this proposed research, previous investigations have determined relationships between achievement and other variables. The inclusion in the linear equation of additional factors suggested through previous research could reduce the unexplained variance of achievement and might reduce the bias in estimates of the parameters of the model. Much of the variance in pupil achievement has a demonstrated relationship with the factors of pupil ability (IQ), race, and socioeconomic status (SES). Therefore, it is proposed to expand the equation to include these variables.

$$Y = B_o + B_1X_1 + B_2X_2 + B_3X_3 + B_4X_4 + \epsilon$$

where: X_2 is the mean pupil ability (IQ)

X_3 is the percentage of pupils in the nonminority population of a school

X_4 is a socioeconomic variable based on the percentage of children in a school who receive free or reduced-price lunches

B_2, B_3, B_4 are factors representing the direction and relative magnitude of the effect of these variables on Y.

Since regression analysis is correlational and does not detect causality, one would not be able tc assert that either achievement or retention caused the other. A rejection of the null hypothesis that there is no relationship between achievement and retention rate does not imply that a change in retention caused a change in achievement. Changes in achievement may precede changes in retention rate, or both factors may be caused by a third factor not examined in the investigation. For example, it might be possible that the state testing program has had an effect on both variables, and that achievement and retention have risen or

fallen with the implementation of state-wide standardized examinations. That is, higher achievement might result from familiarity with the test and with a desire to do well, while increased retention might result from more pressure on teachers held accountable by public awareness of pupil progress. Conversely, as teachers became accustomed to the testing program, they exhibit less concern with retaining low achievers and thus, all pupils achieve better because of the more relaxed atmosphere. Such concerns cannot be addressed with simple correlations.

Path analysis was developed as a method of formulating theory and attaching quantitative estimates to causal effects already hypothesized through previous research or by logical deduction. Although it relies upon the least-squares concept fundamental to regression analysis, path analysis procedures move beyond simple correlations and examine hypothesized chains of causality. Essential to path analysis is the specification of a mathematical model based upon previously espoused theory (Wolfley, 1980).

The variables in the model are arranged in causal order from left to right. Straight arrows signify hypotheses of nonzero causal effects with the arrowhead toward the influenced variable. Curved lines with double arrowheads represent correlations among variables, with causal inference attached. The three variables on the left are collectively referred to as the exogenous variables since their causes originate outside the model. The variables on the right side, excluding S, T, U, V, and W, are referred to as endogenous, since their causes are explicitly included within the model. Variables S, T, U, V, and W are the residual effects upon the endogenous variables.

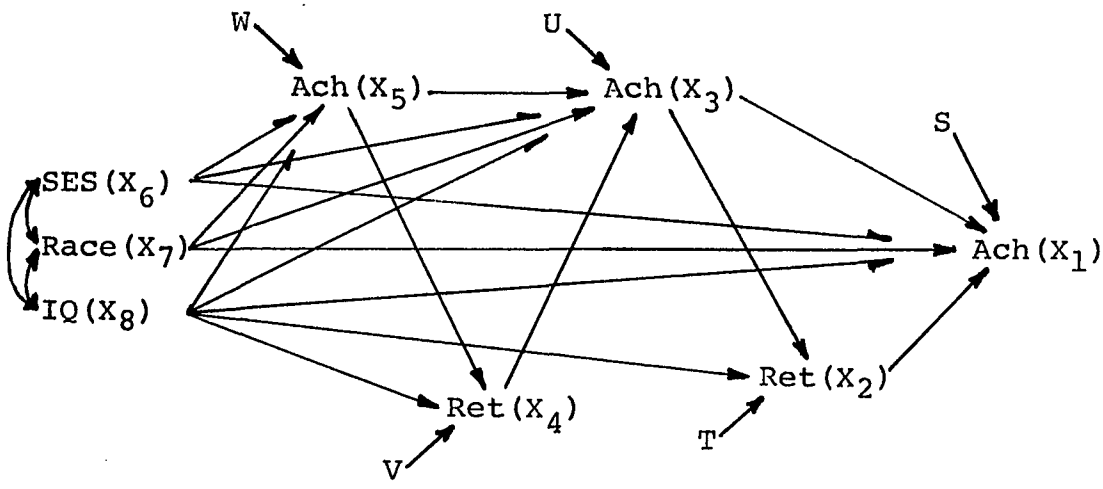


Figure 1. Hypothesized Path Model

According to this model, first-grade achievement (X_5) is caused by SES (X_6), race (X_7) and IQ (X_8). First-grade retention (X_4) is directly caused by first-grade achievement (X_5) and IQ (X_8). Second-grade achievement (X_3) is directly caused by first-grade achievement (X_5) and retention rate (X_4) as well as SES, Race, and IQ. Second-grade retention (X_2) is directly caused by second-grade achievement (X_3) and IQ. Finally, third-grade achievement (X_1) is directly caused by second-grade achievement (X_3) and retention rate (X_2) as well as SES, Race, and IQ. These relationships can be expressed mathematically.

$$X_1 = p_{12}X_2 + p_{13}X_3 + p_{16}X_6 + p_{17}X_7 + p_{18}X_8 + p_{1S}S$$

$$X_2 = p_{23}X_3 + p_{28}X_8 + p_{2T}T$$

$$X_3 = p_{34}X_4 + p_{35}X_5 + p_{36}X_6 + p_{37}X_7 + p_{38}X_8 + p_{3U}U$$

$$X_4 = p_{45}X_5 + p_{48}X_8 + p_{4V}V$$

$$X_5 = p_{56}X_6 + p_{57}X_7 + p_{58}X_8 + p_{5W}W$$

Here the path coefficients (p_{ij}) are standardized regression coefficients where the subscript i refers to the dependent variable. The path coefficient associated with each independent variable in the formula indicates expected change in the standardized dependent variable as a result of a unit change in the standardized independent variable, with the other variables held constant. Further, the application of the path model permits the decomposition of a correlation into its components. Thus, it is possible to estimate what part of each correlation is due to the direct effect of one variable upon another and what part is exerted through an intervening variable within the causal model.

Kerlinger (1973) summarizes the underlying mathematical assumptions for recursive path models as follows:

> (1) The relations among the variables in the model are linear, additive, and causal ....
> (2) The residuals are not correlated among themselves, nor are they correlated with the variables in the system ....
> (3) There is a one way causal flow in the system ....
> (4) The variables are measured on an interval scale ....(p. 309).

Causal effects may be tested with the model by selectively removing paths and examining the effect upon the correlation matrix generated by the complete set of variables. If little discrepancy is noted between the original and the reproduced correlation matrices, it may be concluded that the pattern of the correlations is consistent with the more parsimonious model. If, on the other hand, the reproduced correlation matrix differs from the original matrix after a path is removed from the model, then the theory of this path as a causal link is supported.

The fundamental question under investigation is the relationship between retention rate and subsequent achievement of children in the same school. The proposed model is consistent with both the theorists who argue that retention retards later achievement and those who assert that it promotes greater effort and production. The difference between these theories would be seen in the direction of the relationship between X_1 and X_2, as well as that of X_3 and X_4, as either positive or negative. This study examines not only the supposition of a causal link between retention and achievement, but also the direction of any relationship suggested by the data. Chapter 2 will explore in detail the theoretical and research foundations of this model.

Selection of the Sample

The population studied was the public schools of North Carolina. However, the sample drawn from this population must be considered in light of school organizational patterns, the availability of data, and the stability of the population during the three-year period under investigation. When the 1977 North Carolina Legislature mandated competency testing, it also instituted annual achievement assessment at grades 1, 2, 3, 6, and 9. This testing program has provided a wealth of information concerning pupils' academic progress. It has become possible to examine achievement data for the same group of schools, for children in successive grades, during three consecutive years. Those pupils tested as first-graders in the spring of 1979 were also tested as second-graders in 1980, and as third-graders in 1981. The possibility of examining a single cohort of children in a two-wave model encouraged

the selection of the primary grades for this investigation. In addition, the state has gathered statistics concerning each school's racial and socioeconomic composition and the number of pupils retained in a given grade each year.

However, to properly define the relationships among these factors, school-by-school ability data must also be considered. The state has not universally gathered this information. CTB-McGraw-Hill Testing Service (CTB), which has been awarded the contract for the annual testing program in the primary grades, publishes the Short Form Test of Academic Abilities (SFTAA) as well as the achievement assessment instruments, Prescriptive Reading Inventory (PRI), Diagnostic Mathematics Inventory (DMI), and the California Achievement Tests (CAT). The PRI and DMI are used at grades one and two and the CAT is used at grade three. Because CTB publishes both achievement and ability assessment instruments and can provide predicted achievement scores based on ability, many North Carolina school systems have contracted privately with them for this additional service. Twenty-nine (29) local educational agencies (LEA's) were provided this additional testing in the primary grades in 1981. The 29 systems using the SFTAA contain 379 schools organized with grades one through three on the same campus. This is an average of over 13 schools per system. Nine of these systems account for 256 of the schools, or 68 percent. For economy in data collection, these nine LEA's were selected for this study. Complete data were available for 244 of the 256 schools, thus slightly reducing the final sample.

Of the 145 LEA's existing during the 1980-81 school year, 140 of them had one or more schools with grades one through three organized on

the same campus. Although one unit had 58 schools organized so as to meet this criterion, most systems had far fewer than this. The mean number of school campuses per system including grades one through three was 7.586. The standard deviation around this mean was 7.67. In addition, the median (6) and a mode (4) indicated an asymetrical distribution (Appendix A).

It was necessary to make grade per school, rather than class per school, the sampling unit. The children of one class are frequently dispersed throughout the classes of the subsequent grade, thus precluding the following of individual classes as they progress from grade to grade. However, those pupils promoted from the first grade compose the second grade when adjusted for pupils enrolling in and withdrawing from that group. If the pupils comprising the student body change appreciably from one grade to the next, one can no longer assume that the variables being investigated account for measured differences in achievement. If, however, in- and out-migration is random and normally distributed with respect to the variables being examined, the composition of the sample should remain relatively unchanged. The stability of the sample was examined by comparing the race and SES variables during the first and third years of the study.

The SES variable used in this study is the percentage of each school's enrollment receiving free or reduced-price lunches. Because this is reported by school, rather than by grade, it may not accurately represent the stability of the cohort as it moves from first to third grade. Further, eligibility criteria for free or reduced-price lunches changed during the period of this investigation, thus altering the stability of the SES variable.

The generalizability of the sample to the population of North Carolina public schools was examined by comparing sample and population proportions for racial and SES composition. Since SES proportions were computed from unweighted means they may be less valid than the comparison made with racial proportions. The percentage of white pupils for both the sample and population were computed from raw data.

Availability of Data

Virtually all of the data necessary for this investigation were available from the North Carolina State Department of Public Instruction (SDPI). The Division of Research, SDPI, provided the mean grade achievement per school for grades one, two, and three from the spring testing results of 1979, 1980, and 1981. In addition they provided estimates of racial composition for the same schools, grades, and years. The Information Assistance Service, SDPI, provided retention rates for pupils in the primary grades, for academic years ending in 1979, 1980, and 1981. The Division of Compensatory Education, SDPI, provided SES data consisting of the percentage of pupils in each sampled school receiving free or reduced-price lunches in 1978-79 and 1980-81.

Nine North Carolina public school systems that met the organizational criteria outlined in the Selection of Sample section, above, were contacted and agreed to provide ability data for the schools being studied. IQ data were examined at only one point during the three-year investigation. The information reported was for the most recent ability scores for the cohort finishing the third grade in the spring of 1981. Ability data were gathered at the third-grade level because virtually no North Carolina

school universally assesses first-grade IQ. Although this violates the temporal ordering of IQ as an exogenous variable, it is reasonable to assume that the mean ability of the cohort is the same at the first- and third-grade levels. This assumption was tested by examining the stability of the sample on racial composition during the three-year study. Further, the review of the literature section reports on the stability and reliability of intelligence tests administered at the primary level.

Both the State Department of Public Instruction and the cooperating school systems were assured that no person, class, school, or school system would be identified in any way.

CHAPTER II

REVIEW OF LITERATURE

This chapter is divided into four basic sections. The first of these provides an overview of theories and supporting research concerning the effect of being retained upon individual children. The second section also examines the relationship between retention and subsequent pupil achievement; however, this section looks at promotion rates and achievement means for class groups rather than results for individual children. This is a review of the fundamental research question of this paper. The third section reviews some variables frequently examined in conjunction with retention-achievement research but not included in this investigation. As this section demonstrates a rationale for the selective exclusion of certain factors, the fourth section outlines a rationale for the inclusion of those model variables selected. The fourth section also summarizes the correlational relationships among the factors and thus offers a theoretical framework justifying the model as hypothesized in Chapter I.

Review of Promotion-Retention Theories as Affecting Individual Children

Chase (1968) examined first, second, and third graders reported as basically normal by their teachers but who were immature for their grade. He omitted from this examination children with perceptual dysfunction, or who were of low intelligence, or emotionally disturbed, or who exhibited specific academic problems or poor attendance. Basically, the selected sample was children with perceptual or motor immaturities. Chase felt that it was better to retain children who were immature at the earliest

possible age. Upon examination of these children, it was found that 94 percent showed no ill effects or only slight, short-lived emotional upsets at being retained. The teachers noted six percent as being seriously upset. This was similar to the percentage reported by the parents. Chase concluded that "repeating a grade will engender no negative social or emotional effects in the child whose school failure is based primarily on his immaturity for the grade in which he has been placed" (pp. 176-7). During the following year children with perceptual and motor inabilities developed to a point which more closely approximated the expectations of the teachers than in the previous year during which they were retained.

In a similar study, in the Union Free School District of Valley Stream, New York, Lobdell (1954) reported similar results for children of normal IQ. The success rate for retained elementary pupils was reported to be 69 percent the second time around with only 31 percent of retained pupils making poor progress. It was felt that careful selection of pupils to be retained using specific criteria resulted in the success of this study. Children were retained only if they were of normal intelligence, but showed signs of social, emotional, or personality immaturity. Age and size were also considered as was the attitude of the parents toward the retention of their child. Given these criteria, significant academic progress was reported in subsequent years. In 1960, Worth reexamined the effect of promotion and nonpromotion on the social and emotional development of children. Although much of the pre-1950 research reported adverse affects upon a pupil's self-image, Worth reported positive growth and no adverse effects.

> The findings of this study concerning the affects of promotion and non-promotion on socio-personal development seems to run counter to those reported in the bulk of previous research on this problem. Non-promotion does not appear to have an adverse effect on social-personal development of low achievers. On the contrary, low achievers who are not promoted tend to be rated as high or higher on personality traits and to be accorded the same, and sometimes better, social status than those who are promoted (Worth, 1960, p. 23).

By the early 1970's, some educators were arguing in favor of increased pupil retention. Ellenburg (1972) argued that success is meaningless unless viewed in the perspective of failure. It is failure that permits us to see the true value of success. "We should keep in mind that success cannot evolve without the element of failure present. It is not failure which saves our lives but what is done about failure" (p. 553). He contended that every child should have the right to fail and understand failure as a part of teaching and learning. Pupils should be exposed to the privilege of experiencing selective situations in which they can fail to succeed and yet learn to grow and prosper from that failure.

Scott and Ames (1969) examined a group of retained elementary school children. These pupils had been retained basically because of social and emotional immaturity. An additional factor was assessed prior to retention and following the conclusion of the research. Parents were questioned as to their support or lack of support of the retention of their child based upon the child's immaturity. The pupils acted as their own control and demonstrated significant improvements the second time around. Parents, as well as teachers, were asked to assess the improvements that their children had experienced. It was found that there was a close relationship between the parents' support of retention and the

success of the retention. Although children in general experienced more success when they repeated a grade, those children whose parents had initially supported the retention showed better than average improvement. Thus, the success of retention can be shown to be dependent upon parental attitudes.

Ames (1981) reported that a child's reaction to being retained depends largely upon how his parents view the situation and inform him of his pending retention. She felt strongly that children should be placed in school based upon their behavioral age rather than their chronological or mental age. When children were retained and appropriately placed with children of similar behavioral ages, they demonstrated much higher success. Interviews with four hundred parents of retained children indicated that 90 percent believed the retention was justified and that 87 percent felt that the positive effects of retention outweighed the negative.

Boesel (1960) prepared a doctoral dissertation examining the effect of retention or promotion on two groups of forty-three first grade children. These children were selected from those whose promotions were in doubt from the Toledo, Ohio, schools and were matched into the two groups based upon sex, age, and IQ. In addition, subgroups for boys and girls, low and high IQ, and younger and older first graders were also set up. One year following their retention, all pupils were reexamined to determine the effect of their promotion or retention. The various subgroup examinations showed that children who were younger or of lower IQ profited more from being retained than did older children or children of higher IQ's. However, there were few differences in attitudes either toward

school or toward reading. It was also noted that popularity and social confidence were not seriously affected by a child's retention. It was concluded that children persist in developing in accordance with their natural growth rates since both groups continued to be similar whether promoted or retained. Subsequent follow-up studies on these eighty-six children indicated that promotion or retention makes little difference beyond the initial year in which the child was retained.

In 1977, Bocks reviewed the literature concerning the practice of nonpromotion and concluded that most students do not benefit from being retained. Examined in the study were academic and social factors; however, no statistical data were presented to support this conclusion. Purkey (1978) asserted that school success is based upon positive expectations and positive self-concepts. When children are invited with a warm, cooperative and positive environment, self-concepts can be raised and success elevated. The process of evaluating and grading student achievements as acceptable or unacceptable can destroy a child's self-concept and reduce the likelihood of their success. Godfrey (1971) reported a similar self-concept finding in her work at the North Carolina Advancement School. Retention did not help children catch up and, in fact, retainees were demonstrated to score lower on the Tennessee Self Concept Scale. Poor self-concept may, in fact, interfere with appropriate school achievement.

In 1973, Reiter examined the promotion-retention policies of the Philadelphia school system. He reported that achievement test scores for many repeaters decreased the second year. Supposedly, this was due to the stigma of nonpromotion resulting in poor morale and to the repetition of the same poor teaching methods for a second year. He refuted the claim

that strict promotion policies would help maintain high standards by citing evidence that greater achievement had been demonstrated in systems with more lenient promotion policies than Philadelphia's. In addition, he speculated that larger numbers of retainees lower the work standards for the whole class and that excessive failures indicate already inappropriate standards. He also hypothesized that weak pupils become "teacher pleasers" rather than genuine achievers. Countering these arguments, he felt that automatic promotions (social promotions) are also unsatisfactory for securing pupil achievement. This dilemma, he felt, was not a matter that policy can resolve. Standards must be established but also the child and what is best for him must be considered. Neither promotion for all students nor retention based upon inflexible standards is ideal. In the decision whether to promote or retain, each child must be looked at individually.

In 1959, Worth selected two groups each consisting of sixty-six children who had been matched with respect to sex, IQ, age, achievement scores, and who were in similar schools. One group of low achievers had been promoted to the fourth grade. The other group of sixty-six pupils had been retained in the third grade. Following one year of instruction, these children were reassessed in twelve areas of achievement and eleven areas of social-personal development. In the areas of academic achievement, three scores were significantly higher for the pupils who had been promoted. One subscore was higher for pupils who had not been promoted and eight subscores were not significantly different. When eleven social-personal developmental characteristics were examined, two subscores favored nonpromotion and nine decisions indicated insignificant results.

Worth concluded that although the negative impact of retention may not be as severe as previously hypothesized, there is no indication of a positive effect of retaining a child in grade for an additional year. Gaite (1969) examined the year-end grades for 642 students who had been retained between the eighth and the eleventh grade. He noted that repeating pupils showed significant improvement at the .01 level in some subjects, but not in all. Upon reviewing the previous literature, Gaite noted that Sandin (1944) concluded that mastery of subjects was not assured by repeating a grade and that slow learners were frequently not helped in subsequent years. Further, they noted that high levels of retention adversely affected discipline and that, if promoted, average pupils could catch up. Contradictory research was reported by Steadman (1959) and Lobdell (1954) who found that small groups of carefully selected nonpromoted pupils do improve. In light of these findings, Gaite examined his own results and concluded while there is significant improvement in the grades assigned in the years subsequent to a retention, that the improvements are only slightly better and subject to the "notorious unreliability of teacher assigned marks" (p. 5).

In 1975, Ammons examined the progress of certain students in grades two through five in eight elementary schools in two cities. These pupils had all failed, but in one city, in spite of their academic failure, they had been promoted to the subsequent grade. Pupils were matched on sex, race, IQ, and age. Upon examination of their achievement in the subsequent year, it was found that there were no significant improvements in reading or math achievement and no basic difference in social or emotional adjustment. It was therefore concluded that academic failure

does not in and of itself justify nonpromotion. Of interest was the finding that in spite of formalized policy, teacher decisions tended to be based upon informal policy, often in direct conflict with system-wide standards.

As can be seen from the research reported, the results of retention are often ambiguous or contradictory. In 1975, Jackson conducted the most authoritative review of the literature and summary of the evidence concerning promotion and retention to date. He examined 49 sources reporting original research related to retention and promotion and 54 sources without statistical evidence but discussing related research. His review of the original research found three basic design styles. The first style, which was the most common, compared the results of pupils who were retained under normal school policies with the results for pupils promoted under normal policies. This type of comparison is biased toward indicating that grade promotion has more benefits than grade retention because it compares retained pupils already having difficulties with promoted pupils who are not having difficulties. The second type of design compared the outcomes of retained pupils before and after their retentions. This design is biased towards indicating that pupils benefit from the retention by failing to control for other variables outside of retention itself. By disregarding maturation and other variables affecting the individual from the time of his retention to the time of his future evaluation, one might conclude that all pupils would benefit from retention. The third type of design that Jackson studied was an experimental approach whereby pupils being considered for retention were randomly divided into those retained and those not retained. Only three studies

using this approach were located. The biases reported by Jackson for the first two designs are not necessarily contradictory. This is because the second type of design investigates only the effect of grade retention on low-achieving pupils and the first type of design attempts to compare the effect of grade retention and grade promotion on these pupils. It is possible for promotion to help these same pupils. Jackson concluded "there is no reliable body of evidence indicating that grade retention is more beneficial than grade promotion for students with serious academic or adjustment difficulties" (p. 627). Only three examples of experimental designs were found and the results of these were inconclusive. Since no evidence supports the relationship between retention or promotion and subsequent achievement, an hypothesized correlation of zero between retention and achievement will be used in this study.

Review of Retention-Promotion Theories as Affecting Class Groups

Thus far, the review of literature has examined the affect of promotion or retention on individual children; however, it is also possible to hypothesize the affect upon classroom groups. Kowitz (1961) compared two New York school districts that were similar in composition and size but different in promotion policy. Table 1 compares these systems on five different criteria. These schools were selected for the differences in their retention rates and similarities on other variables. All pupils were examined for achievement over an extended length of time. An examination of pupils' achievement was made each year for ten years comparing achievement with ability at each point. An allometric chart model compared the academic growth in a longitudinal fashion by

examining the percentage of students in each school demonstrating a change in the ratio of achievement age to chronological age.

Table 1
Kowitz Matched Schools

	School A	School B
Membership Grade 10	169	118
Percentage of Grade Failed Before Grade 7	5.92	26.27
Socio-economic Index	3.80	3.33
Composite Achievement Mean	14.44	15.02
Lorge-Thorndike IQ Mean	108	111

Figure 2 demonstrates the graphed achievement in comparison to IQ channels, each channel representing ten IQ points. This permitted charting the number of students crossing one or two channels upward or downward. An upward move would indicate children who were "over-achieving." A downward movement would indicate "under-achievement." Table 2 demonstrates the results achieved at the conclusion of the ten-year study. This was for all pupils in the tenth grade of Schools A and B. Similar results were also obtained when just the pupils who had been retained were examined. The schools were matched for socioeconomic status, composite achievement, and mean IQ, but the slight differences in these factors were not examined or controlled for. While Kowitz concluded that "apparently a policy of high retention can result in a greater proportion of pupils who show an increasing rate of achievement" (p. 441), he also ascribed the differences realized to the total administrative

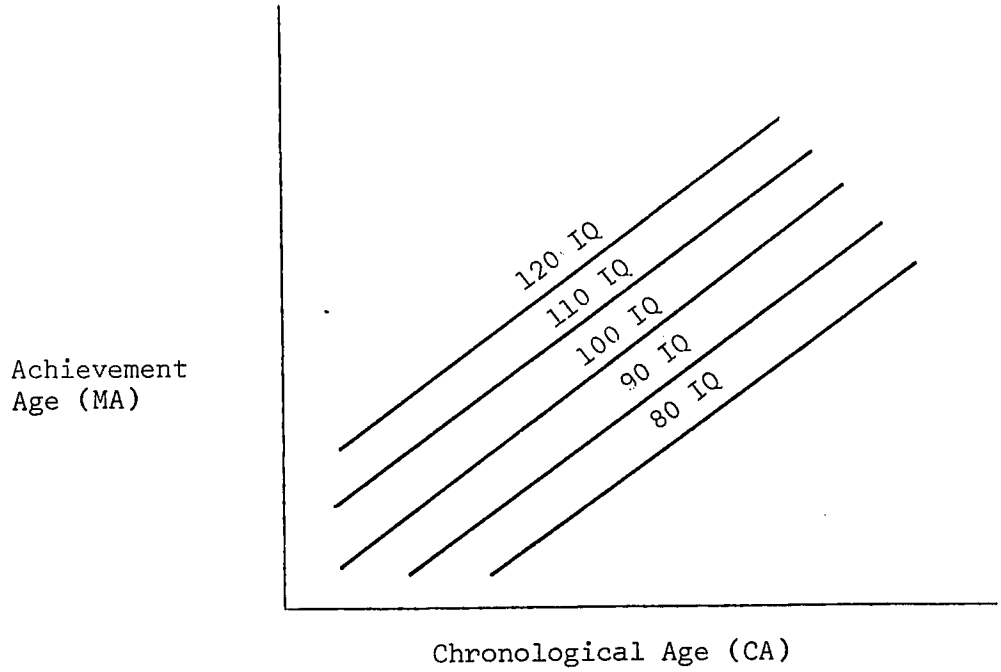


Figure 2. Kowitz Allometric Model

Table 2
Kowitz' Results

Reading Achievement	School A%	School B%
gained 2 channels	2	25 ($p < .01$)
gained 1 channel	27	31
no change	46	31 ($p < .05$)
lose 1 channel	21	13
lose 2 channels	4	0

Math Achievement	School A%	School B%
gained 2 channels	5	21 ($p < .01$)
gained 1 channel	40	42
no change	49	25 ($p < .01$)
lose 1 channel	5	12
lose 2 channels	1	0

climate of the schools, not just retention policy. Because of efforts to maintain high standards of achievement and quality teaching, it may be that greater funds were expended, differences in curriculum existed, and more pressure for achievement was present in one system than in the other. These factors, like SES and IQ, have not been examined. However, it is apparent from the results that school A and school B demonstrated significant differences in achievement when group achievement was compared with retention rates.

The word failure can have two meanings, that is, failure to be promoted or a failure to learn. Ebel (1980) contended that the first of these can be controlled by policy, but not the second; policy cannot effect a child's learning. Ebel insisted that the second quarter of the twentieth century demonstrated a shift in emphasis from failure to learn to conditions of learning, from the product to the process, from subjects to children. Social promotion resulted in concern for the well-being of children in the social setting of the school. Policies of social promotion removed fear of failure as an incentive to work. Social scientists recognized that success is relative, having value only in comparison to failure or as an alternative outcome to human endeavor. Ebel suggested that success cannot be guaranteed, nor failure to learn abolished, since neither is under the control of the teacher. Further, he proposed that there is educational value in failure in that it can teach and can motivate.

Some research has recently indicated improved academic, personal, and social adjustment for groups of pupils who are carefully selected

for repeating a grade under very specific criteria. Owen and Remick (1977) demonstrated overwhelming enthusiasm for the process of retaining non-achieving students. Their enthusiasm is demonstrative of the pendulum swing to a more conservative point of view. They reported the results of a change in policy in the Greenville County, Virginia, schools. Promotion there was made contingent upon demonstrated achievement. Student mastery of skills was required to move from one grade to the next. This policy, which moved away from social promotion and age-based grade assignment toward performance-based assignment, resulted in 1,300 of 3,750 students being retained the first year. While the retention rate dropped under the second and third year to 1,100 and then to 695 students, achievement scores came up dramatically. Students demonstrated more positive attitudes toward testing and reported IQ scores also rose. They noted that the dropout rate declined during the three-year period and that students and teachers appeared to have an increased satisfaction in a job well done. However, this research had significant weaknesses, in that it reported impressions with no statistical analysis. No data were present to demonstrate a relationship between achievement and retention. It was criticized for failure to consider other factors which might affect student achievement besides the retention rate.

Koons (1977) reacted to the Greenville study by citing four possible reasons for the improved achievement found there. First, the fact that the system focused on poor pupil achievement might in itself affect an improvement similar to that of the Hawthorne studies. Secondly, the improved achievement may very well have existed, but at the expense of the retained pupils. Low-achieving pupils, she contended, perform better under a

liberal promotion policy. Some pupils would demonstrate higher achievement because of fear of failure. Third, improved achievement may have resulted from an increased motivation resulting from an awareness that test scores matter. And finally, teachers may have been subconsciously teaching the test to help their students more closely match the teacher's assessment of pupil progress.

Cook (1941) examined the seventh-grade population of two schools that were similar in size, socioeconomic status, and preparation of teachers. He concluded, upon examining the achievement levels of the two groups, that there were no significant differences except in the "over-ageness" of the children from one school. The school with the high rate of retention demonstrated average pupils as being .80 years behind their age mates, while the low ratio school demonstrated average pupils as being .27 years behind their appropriate age placement. While it was reported that the schools were matched on size, SES, and preparation of teachers, there were some differences between the schools. The school reporting a high ratio of retention had more parents who were farmers and slightly skilled laborers, while having fewer parents who were trained professionally, semi-professionally, clerically, managerially, or who were skilled and semi-skilled laborers. In addition, eleven percent of the high ratio school parents were deceased or unreported as opposed to only 4.3 for the low ratio school. No attempt to control these differences was reported. Differences also existed between mean IQ of the high ratio and low ratio schools. That school retaining more pupils reported a mean IQ of 101.58, while the low rate of retention school reported a mean IQ of 106.82 at grade seven. Cook concluded that "the high percentage of over-aged

pupils retained in the upper grades of schools with high standards of promotion reduces the mean intelligence of the classes and lowers significantly the achievement averages of the grades when compared with schools with more lenient standards of promotion" (p. 437).

In 1958, Hall and Demarest reported on the ten-year study of the Phoenix city school system. In 1948, Phoenix changed its academic achievement policy to a continuous promotion policy, finding it desirable to have children stay with their age-mates. Under this policy, children could be retained no more than once in the first three grades and no more than once in grades four through eight. The authors concluded that the data indicated no marked difference in average reading grades or average intelligence scores during the ten-year period. They reported that in 1955-56, 79 percent of the children fell in the normal age range for fourth grade, the grade being examined, as opposed to only 56 percent in 1946-47. However, since the retention rates for each year were not present, a comparison of achievement and retention rates was impossible.

As with research concerning the affect of retention on individual children, many ambiguities still remain as to the effect of retention on groups of children. For the purpose of this study, the anticipated correlation is zero.

Review of Tangential Research Affecting Promotion-Retention Theories

Following an in-depth review of the literature, Purkey (1970) reported a persistent and significant relationship between self-concept and achievement. He found that this relationship was clearer for boys than it was for girls and that it held across racial boundries. Male

under-achievers report lower self-concepts than do females, and blacks report lower self-concepts than do whites. He found that students with low self-concepts rarely perform above average and some who profess to have high self-concepts still do not achieve well. Therefore, he concluded that self-concept was a necessary but not sufficient attribute for success. He concluded that the preponderance of available research indicates that unsuccessful students are likely to hold attitudes about themselves and their abilities which are predominantly negative. They see themselves as less able and less adequate than their more successful classmates. Finally, he concluded that students with negative self-images of ability rarely perform well in school (Purkey, 1970, p. 2). The question remains, however: does poor self-concept cause low achievement or does low achievement cause poor self-concept? The answer to this question is not immediately apparent. It would appear that self-concept can affect academic achievement, both positively and negatively, but also that achievement could result in an alteration of the self-concept. Students carry with them attitudes about themselves and their attributes which affect their performance in school. Yet, scholastic performance can also impact heavily on self-concept (Purkey, 1970, p. 75).

Caplin (1966) examined the relationship between self-concept and achievement in segregated and integrated schools. He found that there is a significant positive relationship between self-concept and academic achievement and also a significant positive relationship between levels of aspiration and achievement. Thus, the higher a student's self-concept and level of aspiration, the higher his reported achievement. Caplin, did not factor out or control for ability or SES. Yet, his findings are significant in terms of a theoretical structure of student achievement.

Intellectually average boys and girls were examined by Farls (1967) to determine the importance of self-concept as a nonintellectual predictor of achievement. He examined boys and girls separately and found, as others have, that high-achieving boys and girls have higher self-concepts and that this relationship was stronger for boys than for girls. He also found that intelligence, as measured by IQ instruments, does not relate to self-concept. Although he found slightly higher mean self-concept scores for girls, he concluded that there was no significant difference in average self-concepts between boys and girls.

In research supported by the Department of Health, Education, and Welfare, Brookover (1969) examined the relationship of self-image to achievement in junior high school students. He found that self-concept was correlated with achievement when IQ is controlled for and that self-concept is also correlated with ability when grade point average is partialed out. He felt this indicated that self-concept is different from intelligence and that self-concept of ability differs from subject to subject. The magnitudes of the correlations between self-concept and achievement with IQ controlled for, and self-concept and ability with grade point average controlled for were similar. This would be expected if a high correlation between achievement and ability were present. Brookover also examined the relationship between SES and self-concept and found the relationship to be positive, but not significant.

In a doctoral dissertation presented at Indiana University, Peters (1968) found results that differed from those of Brookover's research. Peters examined the relationship between self-concept as measured by the Tennessee Self Concept Scale and over- or under-achievement as determined

by discrepancies between predicted grade point average and actual achievement. He examined a wide range of self-concepts, not just academic self-concepts. He found that when achievement is predicted from IQ, and SES is factored out of self-concept, no significant relationship was found between self-concept and over- or under-achievement. He reported this as a complex problem where SES affects self-concept and IQ affects achievement. Self-concept does not affect academic achievement independent of socioeconomic status. Peters' study did not control for sexual differences.

Glick (1969) examined pupil attitudes toward schools, teachers, school work, and peers to see if attitudes affect achievement or if achievement determines attitudes. Using a cross-lagged panel correlation procedure (CLPC), Glick found little evidence to satisfy the claim that either attitudes or achievement was the antecedent of the attitude-achievement relationship. When confronted with significant ambiguity in the CLPC model, Glick examined the "frequency of change in product movement" (p. 3) among the cross-lagged pairs and concluded that

> achievement is more frequently the causal factor but it usually operates to lower the attitude-achievement relationship. When attitudes are the source of influence, it more frequently raises than lowers the attitude-achievement correlation. However attitude influence is relatively infrequent (p. 12).

Glick demonstrated correlations between the Wide Range Attitude Profile (WRAP) and academic achievement ranging between a low of .13 for males to a high of .39 for females. This study has elected not to include a separate variable for self-concept in its path analysis model. It is assumed that attitudes are exogenous variables acting through IQ, Race, and SES, or as a component of uncontrolled variance.

Anxiety and fear of failure have been reported as possible intervening variables in pupil achievement. Sarason (1961) tested the hypotheses that type of instructions and personal degree of anxiety combine and affect performance. An assessment of student anxiety levels was determined and two conditions were established, one providing for non-anxiety-producing instructions and one set of instructions designed to heighten anxieties. It was found that highly anxious students under the threat conditions performed at a lower level than did low- and middle-anxious students. The opposite was found under nonthreat conditions. Already anxious students demonstrated superior performance in test situations with little or no induced anxiety. "The results were interpreted in terms of the interfering response aroused in high anxious individuals under personal threat conditions" (Sarason, 1961, p. 167). Hawkes (1971) examined the relationship between anxiety and achievement while controlling for race, SES, and IQ. He found a significant ($p < .001$) inverse relationship between anxiety and ability ($r = -.31$) as well as between anxiety and achievement ($r = -.42$). IQ and achievement were highly correlated. He concluded that anxiety was a demonstrated motivating force in pupil achievement. Otto and Melbey (1935) attempted to evaluate the threat of failure as a motivating force for pupil achievement. Second- and fifth-grade children who were being considered for possible retention were randomly divided into two equal groups. Half of the children were informed at mid-year that they would not be promoted if they did not improve their work. The remaining half were given no indication of their possible failure. Pupil achievement at the end of the year demonstrated no significant difference between those who were threatened and those who were not.

Fear of failure certainly is closely tied with anxiety and self-concept. Birney, Burdick, and Teevan (1969) discussed at length the defenses against fear of failure. They conclude that the most obvious way to avoid loss in self-estimate, punishment, or loss in social value is to insure the attainment of the standard. Pupils increased the likelihood of achieving what the teacher requires by improving skills through practice and by exerting maximum effort. However, there is no guarantee that the skill will improve and that the goal will be attained. More practice and effort may further reduce self-evaluation if results are not demonstrated. Therefore, a fear of failure is more commonly combated by avoidance behavior. Individuals most frequently first avoid situations where failure will result in punishment or a decrease in peer or self-evaluation. In general, it is doubtful whether additional practice and effort will be exerted by the student who is fearful of failure. His impending failure will certainly be ego-damaging if he must report he tried and failed. It is much easier to report his failure as a lack of effort. If, however, his avoidance of the standard is not possible, a different strategy may be used.

> If the individual cannot escape, either because of high prison walls or the presence of an experimenter or a threat of a truant officer, the question must be answered as to what he will do if faced with an achievement task.
>
> If there is no issue about evaluating the underlying skill the most direct way of avoiding the punishment is through increasing the changes of attainment through practice and effort. There are few achievement situations in which evaluation is not involved and thus the individual has to decide whether devaluation is more or less important to him than the threatened punishment. If the threatened punishment is more fear-provoking, we would expect the person to expend his full effort and take his chances of devaluation. But for this to happen he must be able to see his chances

> of attainment are nil, we expect he will resign himself to the punishment and save the energy of trying (Birney et al., 1969, pp. 220-221)

Schools typically do not offer the elementary student the option of avoiding the fear of failure by removing themselves from the stressful environment of the classroom. Therefore, it is more likely that the pupil will attempt to avoid failure through increased effort and skill development. Grade retention, like subject or test failure, may be seen by the student both as a punishment and as a force devaluing his peer and self-evaluation. If the student perceives a reasonable chance to be promoted to the next grade, fear of failure should motivate practice, effort, and improved academic achievement.

In 1967, Gibby and Gibby investigated the effect of stress induced by academic failure upon a classroom of gifted and talented students. Two classrooms, one acting as a control, were administered three tests and determined to be of similar composition, both in intellectual abilities and personality characteristics. On the first day of the experiment, all students were administered an English grammar test. On the third day of the experiment, both groups were administered the Gibby Intelligence Rating Schedule (IRS) and Fluency Test. On day five the word fluency test and IRS were repeated for both groups, but prior to administering the test to the experimental group, each child was handed a slip of paper reporting his grade on the English test taken on day one as a failure. As a result of the failing grade reported to the students just prior to their taking the test on the fifth day of the experiment, the experimental group was found to have a significant decrease in word

fluency as compared to the control group ($p < .001$). Gibby and Gibby concluded that as a result of the stress of the failure situation, the children regarded themselves less highly and demonstrated a decrement in intellectual productivity.

Cofer and Appley (1964) summarized the work of McClellan, Atkinson, Clark, and Lowell (1953) and reported that motivation is the result of a discrepancy between expectation and perception. One's expectations, referred to as "adaptation level" (AL), foster no motivation when the perception of the environment is as anticipated. Minor discrepancies either positive or negative yield positive motivation. Major discrepancies from what is expected, however, yield negative motivational results. The child who is being rated as he expects is motivated neither to try harder nor less hard. Minor failures, like minor successes, yield positive motivational results balancing the equilibrium between expectation and achievement. However, major positive or negative differences between one's anticipated achievement and what he perceives as achieved produce negative motivational results. This might be illustrated by the gifted child who unexpectedly achieves higher than he anticipated and resultedly demonstrates less motivation. This research might explain in part Gibby's results. Gibby randomly failed some children and subsequently reported a significant decrease in their motivation and production. However, Haber's (1958) butterfly curve would predict just such a relationship. The positive or negative motivational differences result not from the direction of the discrepancy but rather from the magnitude of the discrepancy. Although Haber's study dealt with temporal perception and preferences, his results suggest the hypothetical relationship

proposed between adaptation level, academic achievement, and motivation. It is the size of the discrepancy, not the direction of the discrepancy that determines motivation (Figure 3). Thus failure could be a positive motivator and failure in general is not a negative motivator.

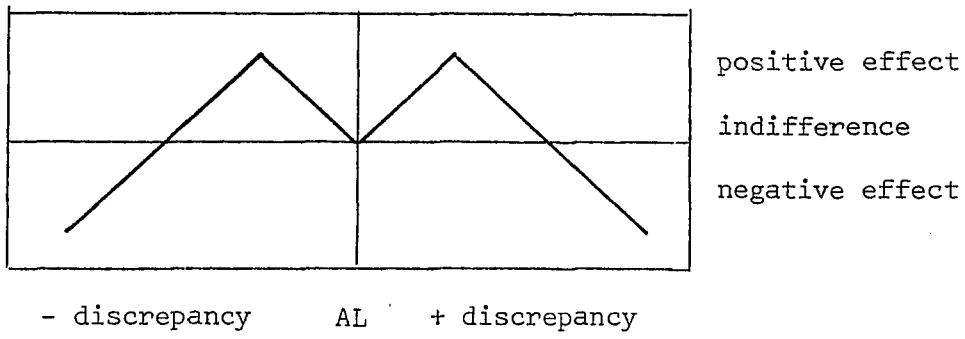


Figure 3. Haber's Butterfly Curve

This research permits some theoretical compromise on the effect of failure on academic achievement. The relationship is far more complex than as hypothesized by parties of either extreme. Neither those arguing that fear of failure is always a positive motivating factor, nor those arguing that failure damages the self-concept and, therefore, is a negative motivator, are entirely correct.

Cofer and Appley (1964) suggest that many conditions may give rise to psychological stress and that any stimulus, no matter how inane, may at some times act as a stress factor but that no stimuli, except possibly life-threatening situations, will universally be stressors. They include fear of failure as one of the seven categories classified as factors giving rise to stress. When either an impossible task is

set or false standards are introduced, the resultant real, contrived, or anticipated failure acts as a psychological stressor. But how the individual deals with that stress is a personal matter. Psychological stress factors may contribute or detract from an individual's orientation. They may result in productive or nonproductive behavior or possibly contribute ambiguous stimulation. Stress is certainly in the eye of the perceiver. The response to stress, like the response to anxiety or a fear of failure, is a personal one. Highly anxious subjects tend to respond to stress in ways that interfere with performance. Yet low-anxiety subjects tend to "energize" in stressful situations and become more task oriented. Because stress and anxiety can either increase or inhibit achievement, they have not been included as model factors in this study.

In 1978, Craig reported on the practice of nonpromotion throughout the state of North Carolina for the years from 1973-1976. He found no relationship between daily absentee rate and promotion rate, no relationship between minority enrollment and promotion rate, no relationship between differing geographic locations or schools of urban or rural composition and promotion rate, nor any relationship between grade grouping patterns and promotion rate. However, a relationship was found between the size of the school systems' enrollment and promotion rate with larger systems demonstrating more retentions. These variables will not be investigated again; however, it is hoped that the results of the current investigation will generalize to the population of North Carolina in light of the work by Craig.

Correlation among Variables to be used in the Model

The following section includes previous research reporting relationships among the model variables. The study currently under consideration uses grade per school as the unit of investigation. However, several of the following citations report correlations for data collected on individuals.

Individual-pupil data are not the same as data reported on classroom or whole-school groups. Mean scores per grade, per school are aggregates of their teachers, classrooms, and pupils. While the effects of education exist for both aggregate and individual data, comparisons made between classrooms or schools are not the same as comparison within these groups (Burstein, 1978, p. 6). Although within-group and between-group scores are not comparable, the reporting of correlations on individuals has been included where comparisons of group mean scores were unavailable or inadequate. Each source has been identified as a comparison of individuals or groups. Expected-correlation matrices at the end of this section summarize the research reported here and separate it according to its source as either aggregate or individual.

Achievement as correlated with retention. The works of Jackson (1975) and Kowitz (1961) as reported in the first and second sections of this chapter indicate the ambiguity surrounding previous attempts to correlate achievement and retention. In general, the most authoritative works indicate significant design problems in attempting to correlate these two variables. Frequently, important variables also correlated with achievement or retention have been omitted or as in the works reported by

Jackson, design faults in the experiments led to unreliable results. Results have indicated both positive and negative correlations between retention and achievement. However, for purposes of this research it will be assumed that there is a correlation of zero between these factors. Jackson (1975) found no consistent correlation when examining the relationship for individuals between achievement and retention in grade. Coffield (1954) also found correlations of zero when examining school means on these two variables.

Achievement as correlated with intelligence. Brookover (1969), whose work with self-concept and achievement was reported earlier in this chapter, also examined correlations between IQ and grade point average (QPA). He found significant correlations for both male and female subjects between IQ and QPA (.61 and .65 respectively). When he factored out social class variables, he found that the correlations were affected, but were still significant. The correlation for males between IQ and QPA with social class controlled was .58, and for females it was .61. All of these correlations were found to be significant. Hawkes and Furst (1971), whose research on anxiety and achievement has already been discussed, also investigated IQ as a correlate of achievement, finding highly significant positive correlations ($p < .001$). They found correlations as low as .78 and as high as .85 between IQ and achievement for various racial and socioeconomic subgroups.

Knief and Stroud (1950) found similar results when they compared achievement with IQ and socioeconomic status. They found correlations between the Lorge-Thorndike Intelligence Test and the Iowa Test of Basic Skills varying between .683 for children of low socioeconomic status and

.839 for children measuring high on socioeconomic status, as indicated by the Warner Index of Status Characteristics. Baker, Schutz, and Hinze (1961) echoed these results, reporting correlations between ability and achievement varying between .46 and .82 for individuals of low and high socioeconomic status. They hypothesized that beyond a certain minimum level of ability, correlations between ability and achievement may be due to uncontrolled status variables. Arena (1970) randomly selected 52 fourth graders and assessed them on levels of intelligence, academic achievement, and social maturity, questioning whether intelligence scores become better predictors of academic achievement when social maturity scores are added to the regression equation. Using the California Test of Mental Maturity, the California Achievement Test Battery, and the Vineland Social Maturity Scale, she found a significant relationship between mental age and academic achievement ($r = .791$). While variance in ability accounted for 55 percent of the variance in achievement, the inclusion in the regression equation of social maturity increased the overall R^2 by only .15 percent. It was concluded that the social maturity scale and mental maturity scale measured many common traits. "Mental maturity is probably the best single predictor of academic achievement available at the present time, although research has shown that it alone does not account for all the variance in academic achievement" (Arena, 1970, p. 21). In surveying the literature concerning factors affecting over- and under-achievement in young school-age children, Asbury (1974) examined pupil scores from grades two through seven and found that achievement correlated with intelligence averaging around .65. He also found that the ocrrelations were better for rural children than for urban children (.67 versus .57) but found

no difference in the effectiveness of prediction for sexual difference in spite of higher failure rates for boys.

Achievement as correlated with race. Innes and Cormier (1973) examined factors which would predict achievement scores based upon non-school factors. Their research was an effort to hold constant the socio-economic factors affecting school achievement to be able to make valid comparisons between schools and to explain the variance between schools on a more equitable basis. They found that the percent of population which is non-minority correlates with measured achievement ($r = .34$). In essence, this demonstrated a correlation of race with achievement. Wise (1976) found similar results in developing a model to predict test scores on the California Achievement Test. He began using socioeconomic variables as a primary source of prediction and found that the addition of racial indicators added considerably to the prediction of average student scores on the achievement test. Fort, Watts, and Lesser (1969) reported research evidence that explains, in part, these differences in racial achievement. They found that different ethnic groups, Chinese, Jewish, Negro, Puerto Rican, display different patterns of mental abilities when assessed by the Diverse Mental Abilities test. This test assesses not only verbal and numerical abilities, but also examines spatial and reasoning tasks. Each group showed unique patterns of strengths and weaknesses. They felt that the differences were due to early family experiences and occupational differences among the ethnic groups. They suggested that achievement tests tend to focus on verbal and numerical tasks and not on spatial and reasoning abilities. Therefore, there should be some difference between ethnic groups on achievement test

results even if there is no overall difference in ability.

Achievement as correlated with socioeconomic status. In predicting achievement test scores for pupils in grades one, two, and three, Wise (1976) used a measure of SES. He selected the variable used to determine eligibility for Title I funding, that is, the percentage of pupils qualified for free and reduced priced lunches, which is based upon a percentage of the population below the federal poverty guideline. This, in addition to average parental education, was correlated with achievement test scores for elementary school pupils. He found that the power to predict reading test scores based upon these variables is not improved substantially by adding more refined measures of education and income. The addition of the percentage of the population which was very poor, or the percentage of the population that was college educated, made negligible differences in the prediction based upon family income and parents' education. Innes and Cormier (1973) found similar results and reported a high correlation ($r = .70$) between measured achievement and socioeconomic status. In their examination of the relationship between achievement, SES, race, desire for education, sex, and family structure, they found that half of the variance in actual achievement can be accounted for by socioeconomic variables. Knief and Stroud (1950) computed intercorrelations among intelligence, achievement, and social variables for 344 fourth-grade students in a midwestern town. Although they found the highest correlation between achievement and ability, they also found significant correlation between SES and achievement and between SES and ability. A correlation of .34 was reported for the 344 pupils between the Warner Index of Status Characteristics and the Iowa Test of Basic Skills.

In an investigation conducted for the Institute of Research on Poverty, Conlisk (1968) examined the affect of demographic variables on the rate of pupil progress. He examined pupil age, race, sex, rural-urban status, education of parents, and parental income and found these variables, especially parental education, to be effective in predicting relative pupil progress. However, he concluded that since these variables are almost completely outside the control of the child, they are to some extent a measure of a lack of equal opportunity. Deutsch, Kate, and Jensen (1968) examined the relationship between socioeconomic status and achievement and found results that were similar to those already reported. They found that social class had a correlation of .44 with reading achievement on the Gates Reading Test ($p < .01$) In 1967 Burkhead, Fox, and Holland examined relationships among inputs, outputs, status, and process variables in large city high schools. They screened out those variables which demonstrated little or no relationship to the others. In general, they found "the lack of variation among school operating characteristics make it difficult to identify factors that produce important differences in school output" (p. 40). These process variables like input characteristics (teacher-principal variables) contributed far less to student outputs than did the status characteristics examined.

Initially, the authors decided to use six status variables, but upon examining the correlation matrix, reduced these to just median family income as the sole status variable because it correlated highly with outputs and summarized the other SES variables well. The authors noted that SES variables are the most important out-of-school variables in predicting a pupil's success and that they are more important than in-school variables.

Teacher experience is important for reading scores, but teachers in low-income schools are younger and have substantially less experience; therefore, there is a correlation between teacher experience and school SES level. Similar to these results found in Chicago were the results found in Atlanta schools. The major determinants of school performance in Atlanta were factors external to the school itself, such as family income and housing conditions. The minor differences in current expenditures had little effect on school outputs. Teacher experience and degrees, as measured by salary, do relate to output but not in a statistically significant fashion (Burkhead et al., 1967, p. 72).

> The most important finding of this study is that variations in educational outcomes in large-city high schools, measured in terms of test scores, are almost wholly conditioned by the socio-economic environment of the neighborhood. The income class of the neighborhood, housing conditions, occupation of parents, ethnic status - these are the important determinants of variations in education outcomes (Burkhead et al., 1967, p. 88).

As with previous sources, these researchers found significant correlations between reading comprehension as a measure of achievement and median family income as a measure of SES ($r = .50$) (p. 79).

Retention as correlated with intelligence. In 1941, Cook examined the effect of the maintenance of high standards of promotion by looking at two schools that were similar in size, SES, and preparation of teachers, but different in terms of the percentage of pupils being retained. Examining at the seventh-grade level, Cook found significant differences between the percentage of pupils who were over age and the mean IQ for the two schools. It was assumed that the increased number of pupils who were over age was

because of a higher rate of retention. The school with the higher rate of retention demonstrated a mean IQ for seventh-grade pupils of 101.58, while the school with the low ratio of retained pupils had a mean IQ of 106.82. This was significant. Although no correlation between IQ and retention was reported, Cook concluded that "the high percentage of over age pupils retained in the upper grades of schools with high standards of promotion reduced the mean intelligence of the classes and lowers significantly the achievement averages of the grades when compared with schools with more lenient standards of promotion" (Cook, 1941, p. 437). While an inverse relationship of unknown magnitude between retention and ability was demonstrated, it is just as logical to assume that the increase in retention rate was due to a decrease in ability as it is to assume the decrease in ability resulted from an increase in retention rate. Clark (1962) reviewed these and similar results of other research and found no significant change in either reading achievement or IQ as a result of a decrease in the number of pupils being retained (p. 56).

Retention as correlated with race. In preparing his 1962 dissertation, Clark reviewed retention rates for North Carolina from 1934 through the early sixties. In 1934, he found a difference in the retention rates in black and white pupils. At that time, 83.3 percent of white pupils were promoted each year as opposed to only 70.9 percent of black pupils. Although these figures decreased during the subsequent years, there were still differences between retention rates for the different races. White schools were still promoting more white pupils than blacks, but as integration approached during the late sixties, the differences were beginning to disappear (Clark, 1962, p. 190). For his 1978 doctoral dissertation

Craig again assessed retention conditions across North Carolina. Among those variables compared with retention rates were the percentages of minority population within each school system. He found no significant differences in retention rates as a result of differences in the percentages of minority population.

Retention as correlated with socioeconomic status and urbanism of schools. Rodgers (1970) examined retention rates as affected by socio-economic status. Although he reported no correlation, he did suggest a hypothetical relationship between these two variables. He predicted that there would be high retention rates in rural areas, very high retention rates in urban centers, but low rates of retention in those suburban areas with high SES. Although Rogers did not investigate these hypotheses, Craig did in 1978. Again, in examining North Carolina school systems and their promotion and retention rates, Craig found no relationship between nonpromotion rates and geographic location nor any difference in promotion rate as related to rural or urban centers. He did, however, find differences in promotion rate based upon size of the school system's student enrollment. A further indication that no relationship exists between promotion rate and socioeconomic status is that Craig found no relationship between promotion rate and funding level through either federal or local governments (1978, p. 43). Federal allocations for Title I Compensatory Education are based upon the socioeconomic status of the school system with higher funding going to systems of lower economic status. Since no relationship between federal funding levels and promotion and retention rates were found, it could be concluded that there exists no relationship between promotion-retention rates and socioeconomic status.

IQ as correlated with race. Burkhead et al. (1967) investigated the inputs and outputs of large city school systems and found significant correlations ($p < .01$) between intelligence and racial factors. They found an inverse correlation between eighth grade IQ and the percentage of the population which was nonwhite equal to -.90. These figures found in Atlanta were echoed by similar results found in Chicago, where the inverse relationship between these two variables was found to be -.86. Backman (1972), in examining the patterns of mental abilities within ethnic groups found that the combined main affects of race, sex, and SES accounted for 90 percent of the total variance in measured ability. Sexual difference accounted for the majority of this (69 percent) with racial difference accounting for 13 percent and SES accounting for only two percent of the total variance found in measured abilities. Ethnic groups for purposes of this study were defined sociologically and not biologically. An R^2 of 13 percent would be expected if the correlation between race and IQ were approximately .36.

Semler and Iscoe (1966) examined the structure of intelligence in black and white children by using the Wechsler Intelligence Scale for Children (WISC) and the Progressive Matrices (PM). The authors reported that although black children function effectively in their environment, reported IQ scores are often lower for them than for white children of the same age. This suggests "that the mental abilities estimates provided by some instruments may be inappropriate for individuals, who, for whatever reason, have need of a functional intelligence quite different from that needed by the standardization population" (p. 327). The authors hypothesized that the WISC verbal scales contained factor

loadings reflecting previous experience in education and that the PM was designed to be less susceptible to the influences of previous learning. Their results tended to confirm these findings. On the WISC they found significant differences ($p < .001$) in favor of white children at all age levels. However, the PM scores showed significant differences ($p < .05$) in intelligence favoring white children only at age seven and not at ages eight and nine. These findings support the authors' contention that differences in measured IQ are in part the result of psychocultural factors. Fort et al. (1979) reported similar findings when using the Diverse Mental Abilities test. They noticed differences in mental abilities between Chinese, Jewish, Black, and Puerto Rican children and hypothesized that the differences were due to early family experiences and occupational differences among the ethnic groups. Using twin comparisons Scarr-Salapatek (1971) determined a decrease in the variability in IQ for children of poverty families and minority races when compared with a randomized population. She reported lower scores on ability for both socially disadvantaged and minority populations and considered this to be a function of the decreased variability of IQ scores in children from disadvantaged homes since these results were independent of race. She noted that the black population she investigated had a higher proportion of disadvantaged persons and therefore a decrease in variability in black scores as well as for low SES scores. She suggested an interaction-interdependence between IQ, race, and SES. In addition, she proposed an hypothesis of cultural difference to account for that portion of interracial IQ differences not attributed to SES differences. Thus, high correlations between race and SES would artificially inflate correlation between IQ and SES.

IQ as correlated with socioeconomic status. As indicated in the previous section, it is quite likely that IQ, race, and socioeconomic status are all closely related one with the other. Table 3 is taken

Table 3
Aptitude as Correlated with Achievement by Scarr-Salapatek

	Blacks	Whites
Low SES	.67	.58
Middle SES	.70	.77
High SES	.71	.72

from Scarr-Salapatek's work and indicates the close correlation between these three variables. Burkhead et al. (1967), in their investigations in Atlanta and Chicago, found results for the correlations between IQ and SES similar to those found for the correlations between IQ and race. In Atlanta, they found eighth-grade IQ correlated highly ($r = .81$) with median family income. Similar results were found in Chicago where IQ correlated with SES ($r = .92$) for ninth-grade ability and median family income. In 1961, Baker et al. noted a triangular relationship among SES, ability, and achievement with both IQ and SES affecting achievement and IQ and SES also correlated. They noted a correlation of .86 between ability and SES for their sampling of eighth-grade students and noted that ability was correlated with achievement from .46 to .82 depending on whether SES was controlled for. They concluded that beyond a certain minimum level of ability, correlations between ability and achievement may be due to uncontrolled status variables. Deutsch et al. (1968) reported correlations between SES and IQ with somewhat less magnitude but also

significant at the .01 level. They found SES to be correlated with IQ as measured by Lorge-Thorndike instrument at .37 and SES correlated with the vocabulary section of the WISC at .49 (p. 99). Knief and Stroud (1950) had early results similar to this. In comparing the Warner Index of Status Characteristics and the Lorge-Thorndike Intelligence Test, they found correlations between .304 and .323 for the various subsections of the Lorge-Thorndike.

Race as correlated with socioeconomic status. The correlation matrix for Innes and Cormier's (1973) work indicates a correlation between SES and race of very low magnitude ($r = -.01$). Higher correlations were found by Wise in developing his model for predicting test scores in California. Wise found negative correlations between black students and socioeconomic status ($r = -.12$) and with Spanish students and socioeconomic status ($r = -.38$). When Burkhead et al. (1967) examined median family income per school with percentage of nonwhite population for each school in Atlanta and Chicago, they found much stronger correlations. SES correlated with race in Atlanta at -.58 and in Chicago at -.79. These figures and those from Wise were all significant at the .01 level. Innes and Cormier's correlations were not significant.

Achievement as correlated with achievement. The model proposed for this investigation requires a correlation component for achievement measured over a one-year period of time. Glick (1969), using a cross-lagged panel correlation procedure, found that achievement correlated from the first year of his investigation to the second year of his investigation at .87 for boys and .93 for girls. For purposes of this investiga-

tion, the correlation of achievement with itself over a period of one year will be assumed to be .90. Appendix B summarized the preceding correlations and sources.

Intelligence as correlated with intelligence. An obvious design fault with the model as proposed in Chapter I is the temporal ordering of intelligence (IQ) data. Because IQ is an exogenous variable, the collection of intelligence data should precede the collection of achievement and retention data. However, few North Carolina school systems collect IQ data prior to the third grade. This is due, in part, to the decreased reliability and validity of tests administered in the early primary grades. Buros (1978) reported that the SFTAA is "highly reliable and reasonably valid, except for the nonverbal test in grades 1-2" (p. 287).

These conclusions are consistent with other sources concerning the reliability of IQ tests in the primary grades. Honzik et al. (1948) found that test-retest reliability for intelligence assessment was only .59 from age four to seven but increased to .78 when examined for the years seven to ten. They concluded that "group prediction is good over short age periods, and mental test scores become increasingly predictive after the preschool years" (p. 315).

Bayley (1949) found similar results when examining the stability of IQ measures between the ages of three months and eighteen years. She found correlations for class mean IQ scores for ages seven to nine equal to .79. However, the class mean score correlations increased to .90 when examined for ages nine to eleven (p. 184).

These findings support the reasonableness of gathering IQ data at the third-grade rather than the first-grade level. Although the temporal

order of the model is violated, the expected reliability of the data is increased.

CHAPTER 3

METHODOLOGY

Source of Data and Data Reduction

Data from 244 schools representing nine North Carolina school systems were arranged and keypunched as fifteen variables. Two additional variables were computed. The Division of Research, State Department of Public Instruction, provided achievement results in math and reading for three consecutive years. The Division of Compensatory Education, State Department of Public Instruction, provided socioeconomic variables in the form of the percentage of pupils receiving free or reduced-price lunches in 1979 and 1981, and the Information Assistance Service, State Department of Public Instruction, provided the percentage of pupils retained in first and second grades in 1979-1980. The nine North Carolina school systems represented in the study provided IQ data for the third grade in 1981. Racial composition data were received from the State Department of Public Instruction. Data were provided in the form of the total pupil population and the total white pupil population in each school. The percentage of white population was computed as part of the analysis.

All data were submitted to the Statistical Package for the Social Sciences (SPSS) condescriptive program. The established means, standard deviations, and ranges were surveyed for reasonableness and indications of faulty data. Two keypunch errors were discovered and corrected. The condescriptive program was then rerun Using the SPSS Pearson product-moment correlation program, a correlation matrix was generated for the corrected data.

The corrected raw data were analyzed using the SPSS regression program. Each variable in the reading achievement model was regressed on all variables that temporally preceded it. A regression analysis was performed with mean first-grade reading as the dependent variable and SES, Race, and IQ as independent variables. First-grade retention rates were then regressed on mean first-grade reading achievement, SES, Race, and IQ. Mean second-grade reading achievement was then regressed on first-grade retention rate, mean first-grade reading achievement, SES, Race, and IQ. Second-grade retention rates were then regressed on mean second-grade achievement, first-grade retention, mean first-grade achievement, SES, Race, and IQ. Finally with mean third-grade reading achievement as the dependent variable, a regression analysis was performed using second-grade retention rate, mean second-grade achievement, first-grade retention rate, mean first-grade achievement, SES, Race, and IQ as the independent variables. These five regression equations were then repeated with first-, second-, and third-grade math achievement substituted for first-, second-, and third-grade reading achievement.

Each series of regression analyses generated 25 standardized regression coefficients used in the model as path coefficients. This permitted the delineation of a "just-identified" model in which every possible path was included. It appeared from a visual examination of the two just-identified models that several insignificant paths had been included. Therefore, all standardized regression coefficients (β's) were examined for significance ($\alpha = .05$) using F ratios. The resulting reduced path model for reading achievement (Figure 4) only eleven causal paths while the reduced model for math achievement (Figure 5)

included thirteen significant paths. The exclusion of nonsignificant paths had little effect on the total coefficient of determination (R^2) for the just identified model when compared with more parsimonious models.

Comparison of Total R^2 Values for Just Identified and Reduced Models

Model	Reading	Math
Just identified model	.648	.539
Reduced model	.629	.517

Kerlinger and Pedhazur (1973, p. 318) advocated a "criterion of meaningfulness" by which a path may be considered for removal from a model. Although they reported no set rule for determining meaningfulness, they suggested that paths with coefficients of less than .05 might be treated as meaningless.

Because all the excluded β's were equal to .04 or less, it was assumed that the indirect causal effects that were lost when these direct causal effects were removed were also insignificant.

Decompositions of causal and noncausal effects were performed for each bivariate relationship indicated on each reduced path model. The direct causal effects were taken from the β coefficients used for the path between each pair. The indirect causal effects were computed by summing the multiplicative effects of all possible paths from one variable to the other. In determining each indirect causal effect, it was important to follow the temporal ordering of each variable along an indirect path. The total causal effect of each bivariate relationship was the sum of the direct and the indirect causal effects.

The total correlation reported in the decomposition tables (Table 8 and 9) was the Pearson correlation coefficient taken from the correlation matrix (Table 6). That portion of the total correlation which was not included in the sum of the direct and indirect paths was reported as the noncausal component of a bivariate relationship.

After each regression analysis of the endogenous variables, the resultant R^2 scores were converted into residual path components. Since the R^2 values indicate the total proportion of explained variance of the measure, the unexplained effects were computed as the square root of one minus the appropriate R^2 value.

Finally, the original correlation matrix was reproduced using the path coefficients of the reduced models. The multiplicative effects of path coefficients along all possible paths between each pair of variables were summed. These resulting values were then compared with the original correlations as a method of testing the reasonableness of the reduced path models.

The Reasonableness of the Reduced Models

The reduced path models, Figures 4 and 5, represent assumed causal relationships among the model variables in the models. Please note the renumbering of variables in Figures 4 and 5 from Figure 1. Those paths which have been removed indicate insignificant direct causal relationships between their respective variables. While removal of a path effectively sets the path coefficient to zero, it also implies that only indirect effects may exist between these variables.

Kerlinger and Pedhazur (1973, p. 317) describe a technique for

testing the reasonableness of a more parsimonious model. They conclude that if, after removing some paths, it is possible to reproduce the original correlation matrix, or closely approximate it, the reduced model is justified. The correlation matrix is reconstructed by summing the multiplicative effects of remaining path coefficients over all possible paths between each pair of variables. It is important to note that this analysis does not determine the theoretical relationships among the variables, but rather indicates whether or not the data are consistent with the reduced model.

When the reproduced correlations approximate the original ones, a judgment to accept the more parsimonious model is made. In general, Kerlinger and Pedhazur (1973, p. 318) indicate that discrepancies which approach .05 are acceptable. Table 4 shows the relationships between the original and reproduced correlations for reading achievement. Of the 28 reproduced correlations, three differ from corresponding original correlations by .06 and two by .05. All the remaining reproduced correlations vary by no more than .04 from corresponding original values. Table 5 shows the relationships between the original and reproduced correlations for math achievement. Only two reproduced correlations deviated from their corresponding original correlations by a magnitude of .06 and three by a magnitude of .05. All the remaining reproduced correlations differed from corresponding original correlations by no more than .04. These data support the acceptance of the more parsimonious models.

Stability of Data

The stability of the population is a critical question for this in-

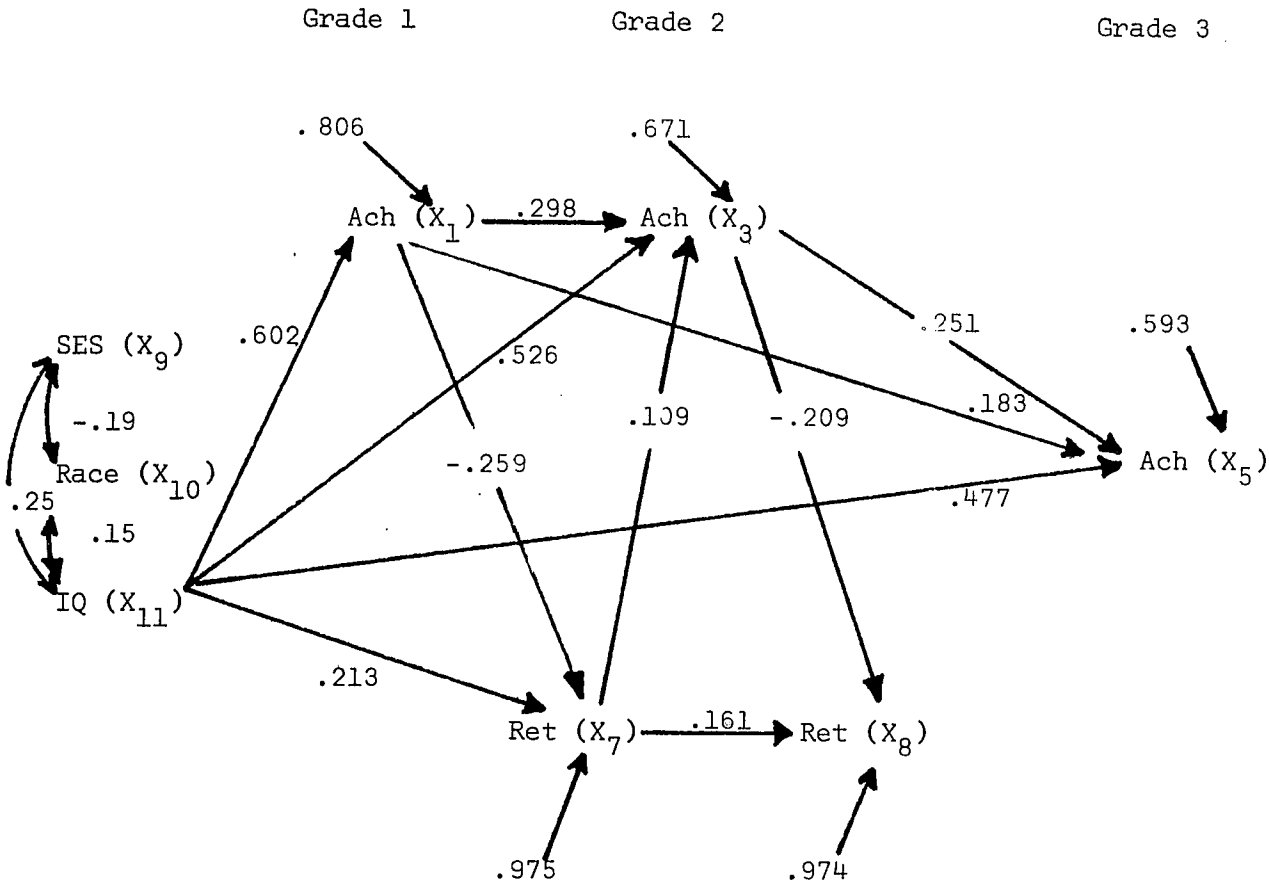


Figure 4. Reduced Path Model for Reading Achievement Means

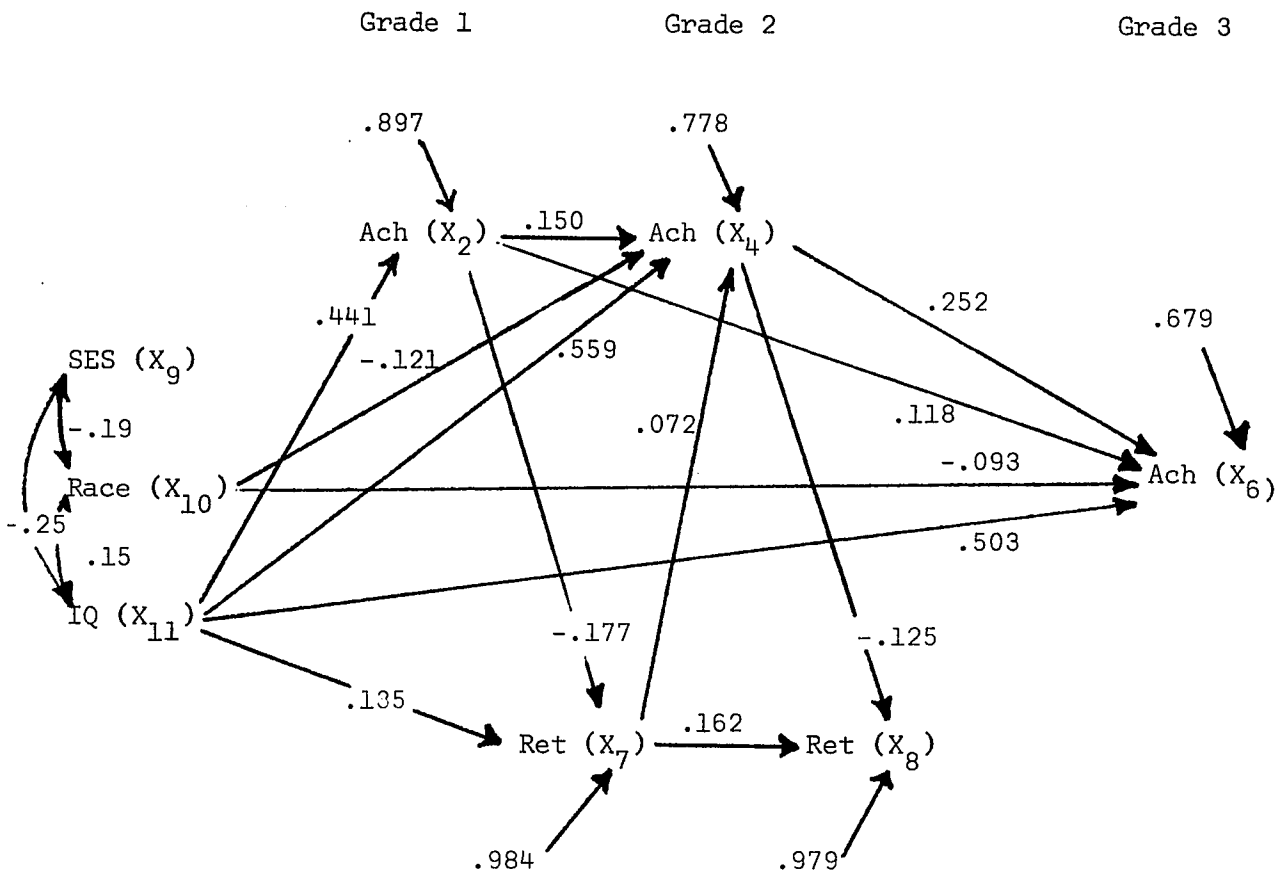


Figure 5. Reduced Path Model for Math Achievement Means

Table 4

Original and Reproduced Correlations for Reading Achievement Means

Code	Measure	X_1	X_3	X_5	X_7	X_8	X_9	X_{10}	X_{11}
X_1	First Grade Mean Reading Achievement	1.00	.61	.62	-.13	-.10	-.16	.09	.60
X_3	Second Grade Mean Reading Achievement	.59	1.00	.64	.15	-.18	-.18	.10	.71
X_5	Third Grade Mean Reading Achievement	.61	.69	1.00	.09	-.04	-.19	.11	.77
X_7	% Retained in First Grade	-.14	.09	.05	1.00	.14	-.01	.01	.06
X_8	% Retained in Second Grade	-.09	-.14	-.07	.14	1.00	.03	-.02	-.12
X_9	% Receiving Free Lunch	-.12	-.15	-.18	.01	.06	1.00	-.19	-.25
X_{10}	% of Population White	.05	.04	.08	-.05	-.07	-.19	1.00	.15
X_{11}	IQ Mean	.59	.70	.76	.05	-.07	-.25	.15	1.00

The original correlations are reported in the lower half of the matrix.
The reproduced correlations are reported in the upper half of the matrix.

Table 5

Original and Reproduced Correlations for Math Achievement Means

Code	Measure	X_2	X_4	X_6	X_7	X_8	X_9	X_{10}	X_{11}
X_2	First Grade Mean Math Achievement	1.00	.37	.38	-.12	-.04	-.11	.07	.44
X_4	Second Grade Mean Math Achievement	.37	1.00	.58	.11	-.10	-.13	-.03	.63
X_6	Third Grade Mean Math Achievement	.42	.59	1.00	.11	-.02	-.15	-.01	.71
X_7	% Retained in First Grade	-.12	.09	.05	1.00	.14	-.01	.00	.06
X_8	% Retained in Second Grade	-.003	-.09	-.02	.14	1.00	.01	-.08	-.06
X_9	% Receiving Free Lunch	-.11	-.07	-.10	.01	.06	1.00	-.19	-.25
X_{10}	% of Population White	.08	-.04	-.03	-.05	-.07	-.19	1.00	.15
X_{11}	IQ Mean	.44	.59	.68	.05	-.07	-.25	.15	1.00

The original correlations are reported in the lower half of the matrix.
The reproduced correlations are reported in the upper half of the matrix.

vestigation. If the population sampled in 1981 is different from that examined in 1979, assumptions concerning the causal linkages between the model variables become invalid. The stability of the sample was determined by examining the consistency of several variables over time.

The racial and socioeconomic composition of the sample was examined when the cohort was in the first and third grade.

The racial composition of each grade per school was reported as the percentage of that group comprising the white population. The correlation between these percentages was computed across schools for the years 1979-1981 and found to be significant ($r = .92$, $p < .01$). A comparison of proportions was performed using the following data:

Racial Composition

Measure	Mean
Mean % of population white in 1979	.676
Mean % of population white in 1981	.682

The equality of these proportions was tested using the statistic

$$z = \frac{p - a}{\sqrt{a(1 - a)/n}} = .200$$

where p was the sample proportion for 1981, a was the sample proportion for 1979, and n equalled 244. With α set at .05, critical limits for z equalled $\pm$ 1.96. Since the computed value of z fell between these limits, the null hypothesis H_o: p = a was not rejected. The sample proportions for the years 1979 and 1981 were assumed to be identical.

The socioeconomic composition of each school was reported as the percent of each school's enrollment receiving free or reduced-price lunches.

In January of 1981 the Department of Education revised the guidelines by which children were approved for lunch assistance and removed all local authority for granting exceptions to the federally approved policy. Therefore, it was not unexpected that SES as measured in 1979 and 1981 was less highly correlated than race. However, the correlation, r = .47, was still significant at the .01 level. A comparison of proportions was performed using the following data:

Socio-Economic Composition

Measure	Mean
Mean % receiving free lunches in 1979	.340
Mean % receiving free lunches in 1981	.353

The equality of these proportions was tested using the statistic

$$z = \frac{p - a}{\sqrt{a(1 - a)/n}} = .448$$

where p was the sample proportion for 1981, a was the proportion for 1979, and n equalled 244. With α set at .05, critical limits for z equalled ± 1.96. Since the computed value of z fell between these limits, the null hypothesis H_o: p = a was not rejected. The sample proportions for the years 1979 and 1981 were assumed to be identical.

In general, these data support the contention that the population sampled in 1981 is the same as that examined in 1979.

Ability to Generalize Data

To generalize the results of this investigation to the population of North Carolina public schools required comparing sample characteristics

with their corresponding population parameters. Population proportions for both racial composition and percentage of pupils receiving free or reduced-price lunches were available from the State Department of Public Instruction.

Of the 84,610 pupils enrolled in the third grade in 1981 state-wide 67.44 percent, or 57,061 were white. Of the 21,544 third-grade pupils who attended the 244 sampled schools, 67.45 percent, or 14,531 were white. The equality of these proportions were tested using the statistic

$$z = \frac{p - a}{\sqrt{a(1 - a)/n}} = .020$$

where p was the sample proportion, a was the population parameter, and n equalled 21,544. Critical limits for z were set at $\pm$ 1.96 (α = .05). Since the computed value for z fell within these limits, the null hypothesis H_o: p = a was not rejected. It was concluded that the sample and population proportions for precentage of white population were identical.

The mean percentage of pupils receiving free or reduced-price lunches in 1981 in the 244 schools sampled was 35.32 percent. This was a composite of unweighted means. State-wide, 39.38 percent of all pupils received free or reduced-price lunches in 1981. The equality of these proportions was tested using the statistic

$$z = \frac{p - a}{\sqrt{a(1 - a)/n}} = 1.33$$

where p was the sample proportion, a was the population parameter, and n equalled 244. Critical limits for z were set at $\pm$1.96 (α = .05). Since the computed value for z fell within these limits the null hypothesis H_o: p = a was not rejected. It was concluded that the sample and population

proportions for percentage of pupils receiving free or reduced-price lunches were the same.

In general, these data support the contention that the selected sample is representative of the population of North Carolina public schools and that the results of this investigation can be generalized to the population.

CHAPTER 4

RESULTS, CONCLUSIONS, AND IMPLICATIONS

Results

Table 6 presents the matrix of correlations between achievement means, retention rates, ability (IQ) means, SES indices, and percentage of white enrollment for the 244 schools. Inspection of this matrix shows that the six measures of achievement were all highly intercorrelated. Math and reading means for the same school grade were more highly correlated than were these variables measured for different school grades. In addition, reading means tended to be more highly correlated than did math means. All correlations between achievement test means were significant at the .01 level.

The percentage of pupils retained in the first grade was significantly correlated ($p < .05$) with achievement at the first-grade level. Retention in the first grade was not significantly correlated with achievement at the second- and third-grade levels. The correlations between retention and achievement means from first to third grade decreased in absolute value from .14 to .05. The correlation between first-grade achievement and retention in the first grade was negative, indicating that higher mean achievement in the first grade tended to be associated with a lower retention rate for the same group of children. The correlations between the percentage of pupils retained in the second grade and all measures of achievement were also negative. However, these correlations were significant only between second-grade reading achievement and retention in the second grade. A significant positive correlation was recorded between first-

Table 6
Intercorrelations Among Model Variables

Measure	1	2	3	4	5	6	7	8	9	10
1. 1st Grade Reading Ach. Mean										
2. 1st Grade Math Ach. Mean	.80*									
3. 2nd Grade Reading Ach. Mean	.59*	.41*								
4. 2nd Grade Math Ach. Mean	.45*	.37*	.77*							
5. 3rd Grade Reading Ach. Mean	.61*	.43*	.69*	.56*						
6. 3rd Grade Math Ach. Mean	.51*	.42*	.58*	.59*	.80*					
7. % Retained in 1st Grade	-.14#	-.12#	.09	.09	.05	.05				
8. % Retained in 2nd Grade	-.09	-.003	-.14#	-.09	-.07	-.02	.14#			
9. % Receiving Free Lunch	-.12#	-.11#	-.15#	-.07	-.18*	-.10	.01	.06		
10. % of Population White	.05	.08	.04	-.04	.08	-.03	-.05	-.07	-.19*	
11. IQ Mean	.59*	.44*	.70*	.59*	.76*	.68*	.05	-.07	-.25*	15*

* $p < .01$ # $p < .05$

grade retention and second-grade retention across schools, indicating that schools that tended to have higher rates of first-grade retention also had higher rates of second-grade retention. However, the correlation coefficient was small (.14).

The percentage of pupils receiving free and reduced-price lunches was negatively correlated with reading achievement means at the first-, second-, and third-grade levels. This correlation increased during the three-year period, from -.12 to -.18. SES was correlated significantly with math achievement only at the first-grade level.

The percentage of the school's enrollment that was white was not correlated with mean achievement in either math or reading at any grade level, indicating that achievement at the primary level is linearly independent of white population percentage.

Mean pupil ability (IQ) was significantly correlated with mean math achievement and mean reading achievement at first-, second-, and third-grade levels. The correlations were higher in reading than in math, and increased from first to second, and from second to third grades for both reading and math. The mean achievement vs. mean ability (IQ) correlations ranged between a low of .44 and a high of .76. These correlations were all significant at the .01 level.

The correlations between SES, race, and IQ were all significant at the .01 level. The percentage of the population which was white correlated positively with mean ability. However, both of these variables had a negative correlation with the percentage of pupils receiving free or reduced-price lunches.

The means and standard deviations for the sample, as recorded in

Table 7, indicate a general increase in the average mean reading and math achievement scores across three grade levels, as well as an increase in the standard deviations of these variables. It is also noted that reading scores tended to be more varied than math scores. In this sample, the average North Carolina elementary school retains 9.4 percent of its first-grade pupils and 5.8 percent of its second-grade pupils. The unweighted mean of school percentage of pupils receiving free lunch equalled approximately one-third. Similar unweighted means indicate that approximately two-thirds of the primary population is white. Mean average school ability (IQ) for the sample was 100.7. The variability of IQ means is, as expected, smaller than the variability of individual IQ scores.

Table 7
Sample Means and Standard Deviations

Code	Measure	Mean	S. D.
X_1	First Grade Reading Achievement	317.8	16.8
X_2	First Grade Math Achievement	338.2	12.7
X_3	Second Grade Reading Achievement	378.2	23.9
X_4	Second Grade Math Achievement	373.8	15.4
X_5	Third Grade Reading Achievement	414.4	21.0
X_6	Third Grade Math Achievement	409.2	14.4
X_7	% Retained in First Grade	9.4	6.0
X_8	% Retained in Second Grade	5.8	4.7
X_9	% Receiving Free Lunches in 1979	33.9	29.5
X_{10}	% of Population White in 1979	67.6	16.9
X_{11}	Mean Ability	100.7	5.0

The Reading Achievement Path Model. The path model as proposed in Chapter 1 included sixteen causal links among the included variables. These data support the existence of nine of those paths. Eight of the anticipated paths were not confirmed, and two unanticipated paths were supported by the analysis.

Ability (IQ) has a strong determining effect upon achievement at first-, second-, and third-grade levels. However, the expected effects of SES and Race upon Achievement were not confirmed. The relationship between either SES, or Race, and Achievement can be explained through their correlation with IQ. No effect of SES or Race on Achievement or Retention was found to exist except as they might act through IQ.

The anticipated relationship between IQ and retention rate was confirmed at the first-grade level. However, no relationship between ability and second-grade retention rate was found. Further, the direct relationship between IQ and first-grade retention was positive. This indicates that those schools with higher mean abilities tend to retain more pupils in the first grade. However, the magnitude of this effect is partially cancelled through negative, indirect effects acting through first-grade achievement. The total correlation between ability and first-grade retention rate was explained through the direct and indirect causal links of the confirmed model.

The relationship between first-grade achievement and first-grade retention rate was found to be inverse. It is logical to assume that those schools exhibiting lower mean achievement would report a higher percentage of pupils being retained. This relationship was also found to hold between second-grade achievement and second-grade retention, although the magnitude

of the path coefficient was only 80 percent of that at the first grade. The general decomposition table (Table 8) indicated a positive non-causal component of .12 between first-grade achievement and retention and .07 between second-grade achievement and retention. While the causal component of the relationship between achievement and retention is negative, the noncausal component of the bivariate relationship is positive. This may indicate that in some schools factors other than achievement are dominant in the decision to promote pupils.

It was expected that first-grade achievement mean would have a direct causal effect on second-grade achievement mean and that second-grade achievement mean would have a direct causal effect on third-grade achievement mean. It was not anticipated that first-grade achievement would have a direct causal effect on third-grade achievement, independent of second-grade achievement. First-grade achievement contributes more toward second-grade achievement than either of these contributes toward third-grade achievement. Because relationships among achievement are positive, success tends to breed success and failure predicts future failures. While first-grade achievement is a significant predictor of second-grade success, it only has 60 percent of the direct predictive effect of ability. Ability continues to be the best single predictor of achievement at the third-grade level; however, the total direct and indirect effects of previous achievement surpass even IQ in predicting third-grade achievement.

The decomposition table (Table 8) indicates comparatively high noncausal components for the bivariate relationships among first-, second-, and third-grade achievement. This implies that one or more variables contributing to achievement have been omitted from the model. Further,

this indicates that the residual components of achievement may not be independent.

Table 8
General Decomposition Table for Reading Achievement Means

Bivariate Relationship	Total Correlation (A)	Causal			Noncausal (E) A - D
		Direct (B)	Indirect (C)	Total (D) B + C	
X_1X_{11}	.59	.60	.0	.60	-.01
X_7X_1	-.14	-.26	.0	-.26	.12
X_7X_{11}	.05	.21	-.16	.05	.0
X_3X_1	.59	.30	-.03	.27	.32
X_3X_7	.09	.11	.0	.11	-.02
X_3X_{11}	.70	.53	.18	.71	-.01
X_8X_3	-.14	-.21	.0	-.21	.07
X_8X_7	.14	.16	-.02	.14	.0
X_5X_3	.69	.25	.0	.25	.44
X_5X_1	.61	.18	.07	.25	.36
X_5X_{11}	.76	.48	.18	.66	.10

The fundamental hypothesis of this research hinges on the relationship of first-grade retention to second-grade achievement and second-grade retention to third-grade achievement. The bivariate relationships X_7X_3 and X_8X_5 address the question of a causal link between retention and subsequent achievement. While the correlation between first-grade retention and second-grade achievement was not significant (.09), the path from X_7 to X_3 was found to contribute to the predictive value of the model. The F ratio

contributing the path coefficient X_7X_3 was significant ($p < .01$). The positive relationship between X_7 and X_3 indicates that those schools having higher retention rates in the first grade are more likely to demonstrate higher mean achievement in the second grade. However, the relative weight of first-grade retention is approximately 36 percent that of first-grade achievement. This model supports the claim that increased first-grade retention has a positive causal effect upon second-grade achievement. It also supports the claim that success in the first grade has a positive causal effect upon second-grade achievement. Further, the relative weights of X_1X_3 and X_7X_3 would indicate that achievement is more likely to cause subsequent achievement than is non-promotion.

Neither the correlation between second-grade retention rate and third-grade achievement nor the path coefficient for these variables indicated a significant relationship. Within the limits of this model, there is no linear relationship between nonpromotion rate and subsequent achievement beyond the first-grade level.

An unexpected relationship was found between first- and second-grade retention rates. A positive path coefficient of .161 supports a significant causal link between first-grade retention rate and second-grade retention rate. Those schools which retain more pupils in the first grade are more likely to retain more in the second grade. This may be the result of specific school policy concerning assignment to a grade or informal but shared practices among a faculty.

The residual coefficients indicated that much of the causal structure of achievement and most of the causal structure of retention have not been included within the model.

The Math Achievement Path Model. The confirmed path model for math achievement was similar to the reading achievement model. Additional paths were discovered between first- and third-grade math achievement and between first- and second-grade retention rates. Further, the expected path between second-grade retention rate and third-grade achievement in math was not confirmed.

SES had no direct effect upon math achievement at grades 1, 2, or 3 as well as no direct effect upon retention rates at either first or second grade. Race, on the other hand, demonstrated a direct effect upon both second- and third-grade achievement. The decomposition of the bivariate relationships $X_{10}X_4$ and $X_{10}X_6$ resulted in total causal effects that were close to zero. While the direct effects of race on achievement were negative, the indirect effects through IQ were positive and of similar magnitude.

In general, the path coefficients for the math achievement model were similar in magnitude to those of the reading achievement model. The strengths of the relationships were consistently less only between first-grade achievement and first-grade retention, first-grade retention and second-grade achievement, and second-grade achievement and second-grade retention. This indicates a weaker relationship between math achievement and promotion-retention decisions than between reading achievement and subsequent assignment to a grade.

The decomposition of the bivariate relationships among first-, second-, and third-grade achievement indicates a large non-causal component. Between 57 percent and 64 percent of the total linear relationship among these variables is non-causal. As with the reading model, this indicates possible omission of causal variables or high correlations among the achievement residuals.

An examination of the decomposition (Table 9) of the bivariate relationships between first-grade achievement and retention (X_7X_2) and second-grade achievement and retention (X_8X_4) reveals some similarities to the reading model. Although the relationship between first-grade achievement and subsequent promotion or retention is inverse, the non-causal component of that relationship is positive. Again, this may

Table 9

General Decomposition Table for Math Achievement Means

Bivariate Relationship	Total Correlation (A)	Causal			Noncausal (E) A - D
		Direct (B)	Indirect (C)	Total (D) B + C	
X_2X_{11}	.44	.44	.0	.44	.0
X_7X_2	-.12	-.18	.0	-.18	.06
X_7X_{11}	.05	.14	-.08	.06	-.01
X_4X_7	.09	.07	.0	.07	.02
X_4X_2	.37	.15	-.01	.14	.23
X_4X_{10}	-.04	-.12	.09	-.03	-.01
X_4X_{11}	.59	.56	.05	.61	-.02
X_8X_4	-.09	-.12	.0	-.12	.03
X_8X_7	.14	.16	-.01	.15	-.01
X_6X_4	.59	.25	.0	.25	.34
X_6X_2	.42	.12	.03	.15	.27
X_6X_{10}	-.03	-.09	.08	-.01	-.02
X_6X_{11}	.68	.50	.19	.69	-.01

indicate that some schools consider factors other than achievement in promoting pupils.

The relationship between first-grade retention and achievement at the second-grade level is positive. However, the magnitude of the relationship is quite low. Although this indicates that schools with higher first-grade retention rates demonstrate higher second-grade math achievement, the relationship should be interpreted only with great caution. Except as noted above, the noncausal components of the bivariate relationships among these variables were close to zero ($\pm$ .02).

The residual components of achievement and retention were slightly larger in the math model than in the reading model. In general, the similarities between the two models indicate that one model is sufficient to describe the relationship among mean math or reading achievement and nonpromotion, SES, Race, and IQ.

Conclusions

The results of this research support the contention that those schools demonstrating a higher rate of retention in the first grade produce somewhat higher mean achievement in the second grade. This relationship holds in the context of previous achievement and ability but is weaker than relationships involving either of these variables. Although the relationship between retention and subsequent achievement is statistically significant, its magnitude is very small.

The confirmed relationship between first-grade retention rate and second-grade achievement is not supported by this research beyond the second-grade level. While retention may provide benefits or harm to

individual children, the achievement effects upon an entire school grade appear to cease with the second grade.

School attitudes toward retention play an important role in determining the percentage of pupils retained in a grade. Schools that retain more pupils in the first grade are likely to retain more pupils in the second grade. This effect is linearly independent of the school's mean achievement.

This research found no significant effect of SES or race upon either mean achievement or retention rate. Any effect of SES or race on these endogenous variables was indirect. Socioeconomic or racial factors affect achievement and retention only through their significant correlations with pupil ability.

Although the causal relationship between mean achievement and retention rate was found to be inverse, the noncausal relationship between these variables was positive. This implies that other factors, such as teacher bias, or school policy, which were not examined in this study, also affect promotion rates. These omitted factors apparently favor promoting rather than retaining pupils.

The effect of achievement upon subsequent achievement is far more powerful than the effect of retention upon subsequent achievement. Retention rate in the first grade had only 36 percent of the effect on second-grade achievement as that of first-grade achievement. This supports those theorists who have argued that success is a better motivator than failure.

Limitations and Suggestions for Future Study. Evidence in the results section of this chapter indicates that the path analysis model used in this

research is lacking in several respects. The noncausal components of the decomposition for several pairs of variables indicate that there may be specification bias in the models. Future research in this area should incorporate measures of learner characteristics as well as teacher or school descriptors in an attempt to reduce specification bias.

A replication of this study using individual rather than aggregate data would permit some of these changes. If individual scores were examined, factors for pupil sex could be included as well as more refined measures of SES or ethnic origin.

This study has not addressed specific teacher characteristics; yet there is evidence in the results suggesting that teacher bias or school policy may affect retention rates. If this study were replicated with individual data, teacher attitudes concerning retention could be assessed. Other teacher variables such as years of experience, degree level, age, or sex may further reduce the residual variance in a model of the relationship between mean achievement and retention rate.

School characteristics have been omitted completely from this research. However, the results demonstrate that schools which retain a higher percentage of pupils in the first grade also retain more pupils in the second grade. Any replication of this work should consider variables to measure school-wide factors. Such factors might include school policy descriptors, the principal's attitude toward retention, the rural or urban location of the school, or the size of the school.

One of the assumptions of recursive path analysis models is the independence of the residual variables. The comparatively high values for the noncausal components of achievement raise questions concerning the

satisfaction of this assumption. It is not unrealistic to assume that achievement means at the first, second, and third grades are reasonably similar. It is also likely that the residuals whould be highly correlated if the model had been incompletely specified. A more completely specified model would not only reduce the residuals but also reduce any intercorrelations among them.

One of the more significant shortcomings of this research is the limited age range which has been investigated. Future researchers may wish to focus on the relationships between achievement and retention during the intermediate or secondary years.

A review of the results, conclusions, and implications indicates to educational policymakers that retention can have a limited, positive effect upon achievement if implemented at the first-grade level. Further, the data support the contention that the results can generalize to the population of North Carolina public schools.

Finally, the path analysis procedures used in conducting this research are judged to adequately meet the needs of educational investigators. While the procedure cannot confirm causal linkages between variables it can support an hypotnesis of causation based on previously espoused theory. To this end path models could contribute significantly to our knowledge concerning the education of children.

BIBLIOGRAPHY

Ames, L. B. Retention in grade can be a step forward. The Education Digest, March 1981, pp. 36-7.

Ammons, J. D. A study of the effects of nonpromotion and promotion as related to achievement and self-concept of elementary school students (Doctoral dissertation, East Texas State University, 1975). Dissertation Abstracts International, 1976, 36, 5011-A. (University Microfilms No. 76-4617)

Arena, T. Social maturity in the prediction of academic achievement. Journal of Educational Research, 1970, 64, 21-22.

Asbury, C. A. Selected factors influencing over- and under-achievement in young school-age children. Review of Educational Research, 1974, 44, 409-28.

Backman, M. E. Patterns of mental abilities, ethnic, socioeconomic, and sex differences. American Educational Research Journal, 1972, 9, 1-12.

Baker, R. L., Schutz, R. E., & Hinze, R. H. The influence of mental ability on achievement when socioeconomic status is controlled. Journal of Experimental Education, 1961, 30, 255-258.

Bayley, N. Consistency and variability in the growth of intelligence from birth to eighteen years. Journal of Genetic Psychology, 1949, 75, 165-196.

Birney, R. C., Burdick, H., & Teevan, R. C. Fear of Failure. Princeton, N. J.: Van Nostrand Co., 1969.

Bocks, W. M. Nonpromotion: A year to grow? Educational Leadership, 1977, 34, 379-382.

Boesel, F. F. Effects of nonpromotion on reading achievement and behavior problem tendencies in the primary grades (Doctoral dissertation, University of Michigan, 1960). Dissertation Abstracts International, 1961, 21, 2191. (University Microfilms No. 60-0839)

Bowles, S. & Levin, H. The determinants of scholastic achievement: An appraisal of some recent evidence. The Journal of Human Resources. 1968, 3, 3-24.

Brookover, W. B. Self-concept of ability and achievement. Michigan State University, 1962. Paper presented at American Education Research Association Convention, Los Angeles, 1969.

Burkhead, J., Fox, T. G. & Holland, J. W. Input and output in large city high schools. Syracuse, N. Y.: Syracuse University Press, 1967.

Buros, O. K. The eighth mental measurements yearbook. Highland Park, New Jersey: The Cryphon Press, 1978.

Burstein, L. Assessing differences between grouped and individual-level regression coefficients: Alternative approaches. Sociological Methods and Research, 1978, 7, 5-28.

Caplin, M. D. The relationship between self-concept and academic achievement (Doctoral dissertation, Columbia University, 1966). Dissertation Abstracts International, 1966, 27, 979A. (University Microlfilms No. 66-10, 284)

Caplin, M. D. Relationship between self-concept and academic achievement. Journal of Experimental Education, 1969, 37, 13-16.

Chase, J. A. A study of the impact of grade retention on primary school children. Journal of Psychology, 1968, 70, 169-177.

Clark, A. O. Pupil promotion practices and policies in the elementary school, with particular reference to North Carolina (Doctoral dissertation, Duke University, 1962). Dissertation Abstracts International, 1963, 24, 1896. (University Microfilms No. 63-7010)

Cofer, C. N., & Appley, M. H. Motivation: Theory and research. New York: John Wiley and Sons, 1964.

Coffield, W. H. A longitudinal study of the effects of nonpromotion on educational achievement in the elementary school (Doctoral dissertation, State University of Iowa, 1954). Dissertation Abstracts International, 1954, 14, 2291-2. (University Microfilms No. 10,200)

Coffield, W. H. & Blommer, P. Effects of nonpromotion on education achievement in the elementary school. Journal of Educational Psychology, 1956, 47, 235-250.

Conlisk, J. Determinants of school enrollment and school performance. Madison: Wisconsin University Press, 1968. (ERIC Documents Reproduction Service No. ED012091)

Cook, W. W. Some effects of the maintenance of high standards of promotion. Elementary School Journal, 1941, 41, 430-437.

Craig, D. H. A study of nonpromotion in North Carolina public school systems from 1973 through 1976 (Doctoral dissertation, University of North Carolina at Greensboro, 1978). Dissertation Abstracts International, 1978, 39, 3253-A. (University Microfilms No. 7824396)

Deutsch, M., Katz, I., Jensen, A. (Eds.). Social class, race, and psychological development. New York: Holt, Rinehart, & Winston, 1968.

Ebel, R. L. The failure of schools without failure. Phi Delta Kappan, 1980, 61, 386-88.

Ellenburg, F. C. Permit pupils to fail! Clearing House, 1972, 46(a), 553-555.

Eysenck, H. J. Race, intelligence, and education. New Society, 1971, 17, no. 455, 1045-7.

Farls, R. J. High and low achievement of intellectually average intermediate grade students related to the self-concept and social approval (Doctoral dissertation, University of Pittsburgh, 1967). Dissertation Abstracts International, 1967, 28, 1205. (University Microfilms No. 07-12, 388).

Feldmesser, R. A. The positive functions of grades. Paper presented at the 55th annual meeting of the American Education Research Association, Washington, D.C., February 4, 1971. (ERIC Document Reproduction Service No. ED 049 704)

Fort, J. G., Watts, J. C., & Lesser, G. S. Cultural background and learning in young children. Phi Delta Kappan, 1969, 50, 386-388.

Gaite, A. J. H. On the validity of nonpromotion as an education procedure. Madison: University of Wisconsin Press, 1969. (ERIC Documents Reproduction Service No. ED046043)

Gibby, R. G., Sr., & Gibby, R. G., Jr. The effects of stress resulting from academic failure. Journal of Clincial Psychology, 1967, 23, 35-37.

Glick, O. The interdependence of sixth graders' school attitudes and academic performance. Paper presented at the Western Psychological Association Convention, Vancouver, British Columbia, 1969. (ERIC Documents Reproduction No. ED035033)

Godfrey, E. The tragedy of failure. North Carolina Education, October 1971, pp. 10-11.

Gold, R. M. Academic achievement declines under pass-fail grading. Journal of Experimental Education, 1971, 39, 17-21.

Haber, R. N. Discrepancy from adaptation level as a source of affect. Journal of Experimental Psychology, 1958, 56, 370-375.

Hall, W. F., & Demarest, R. Effects on achievement scores of a change in promotional policy. Elementary School Journal, 1958, 58, 204-7.

Haney, W. & Madaus, G. Making sense of the competency testing movement. Harvard Educational Review, 1978, 48, 462-84.

Hanson, G. Predictors of achievement: A bibliography. Washington, D. C.: Department of Health, Education, and Welfare, 1973. (ERIC Document Reproduction Service No. ED984629)

Hawkes, T. H., & Furst, N. F. Race, socioeconomic situation, achievement, IQ, and teacher ratings of student's behavior as factors relating to anxiety in upper elementary school children. Sociology of Education, 1971, 44, 333-50.

Heise, D. R. Causal analysis. New York: Wiley and Sons, 1975.

Hess, F. Issues in education: A documented look at seven current topics. New York: East Syracuse - Minoa Schools, 1978. (ERIC Document Reproduction Service No. ED158391)

Honzik, M. P., Macfarlan, J. W., & Allen, L. The stability of mental test performance between two and eighteen years. Journal of Experimental Education, 1948, 17, 309-324.

Innes, T. C., & Cormier, W. H. The prediction of achievement means of schools from non-school factors through criterion scaling. Paper presented at the University Educational Research Association Convention, New Orleans, 1973. (ERIC Documents Reproducation Service No. ED074142)

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

Kenny, D. A. Correlation and causality. New York: Wiley and Sons, 1979.

Kerlinger, F. N., & Pedhazur, E. J. Multiple regression in behavioral research. New York: Holt, Rinehart, and Winston, 1973.

Klecka, W. R., Nie, N. H., and Hull, C. H. SPSS primer. New York: McGraw-Hill, 1975.

Knief, L. M. & Stroud, J. B. Intercorrelations among various intelligence, achievement, and social class scores. Journal of Education Psychology, 1950, 50, 117-120.

Koons, C. L. Nonpromotions: A dead-end road. Phi Delta Kappan, 1977, 58, 701-02.

Kowitz, G. T., & Armstrong, C. M. The effect of promotion policy on academic achievement. Elementary School Journal, 1961, 61, 435-443.

Land, K. C. Principles of path analysis. In E. F. Borgatta & G. W. Bohrnstedt (Eds.), Sociological methodology. San Francisco: Jossey-Bass, 1970.

Lennon, R. T., & Mitchell, B. C. Trends in age-grade relations. School and Society, 1955, 82, 123-125.

Lloyd, D. N. Reading achievement and its relationship to academic performance: Part III Relationships of family background and third-grade performance to sixth-grade reading achievement. Rockville Md.: National Institute of Mental Health, 1972, (ERIC Documents Reproduction Service No. ED064704)

Lobdell, L. O. Results of a nonpromotion policy in one school district. Elementary School Journal, 1954, 54, 333-337.

Mandler, G. & Sarason, S. B. A study of anxiety and learning. Journal of Abnormal Social Psychology, 1952, 47, 166-173.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. The Achievement Motive. New York: Appleton-Century-Crofts, 1953.

Milner, D. C. Three factors of school achievement. Journal of Educational Research, 1967, 60, 370-76.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, & Bert, D. H. SPSS: Statistical Package for the Social Sciences. New York: McGraw-Hill, 1975.

North Carolina State Department of Public Instruction. Statistical Profile: North Carolina Public Schools. Raleigh: North Carolina State Department of Public Instruction, 1975.

North Carolina State Department of Public Instruction. Statistical Profile: North Carolina Public Schools. Raleigh: North Carolina State Department of Public Instruction, 1980.

Otto, H. J., & Melbey, E. O. An attempt to evaluate the threat of failure as a factor in achievement. Elementary School Journal, 1935, 25, 588-96.

Owen, S. A., & Ranick, D. L. The Greenville program: A common sense approach to basics. Phi Delta Kappan, 1977, 58, 531-33, 539.

Peters, D. M. The self-concept as a factor in over- and under-achievement (Doctoral dissertation, Indiana University, 1968). Dissertation Abstracts International, 1968, 29, 1792f. (University Microfilms No. 68-17,289)

Purkey, W. W. Self-concept and school achievement. Englewood Cliffs, N. J.: Prentice-Hall Inc. 1970.

Purkey, W. W. Inviting school success: A self-concept approach to teaching and learning. Belmont, California: Wadsworth, 1978.

Reiter, R. G. The promotion/retention dilemma: What research tells us. Philadelphia: Pa. Office of Research and Evaluation, Philadelphia School District, 1973. (ERIC Documents Reproduction Service No. ED099412)

Rodgers, W. A. Retention as a school policy. The Urban Review, 1970, 4, 29-30.

Roth, R. M. The role of self-concept in achievement. Journal of Experimental Education, 1959, 27, 265-81.

Sandin, A. A. Social and emotional adjustment of regularly promoted and nonpromoted pupils. Child Development Monographs, No. 32. New York: Columbia University, 1944.

Sarason, I. G. The effects of anxiety and threat on the solution of a difficult task. Journal of Abnormal and Social Psychology, 1961, 62, 165-168.

Scarr-Salapatek, S. Race, social class, and IQ. Science, 1971, 174 (4016), 1285-1295.

Scott, B. A., & Ames, L. B. Improved academic, personal, and social adjustment in selected primary school repeaters. Elementary School Journal, 1969, 69, 431-439.

Semler, I. J., & Iscoe, I. Structure of intelligence in Negro and white children. Journal of Educational Psychology, 1966, 57, 326-336.

Steadman, E. R. Fifteen who were not promoted. Elementary School Journal, 1959, 59, 271-276.

Test scores improve in structured setting. Guideposts, 1973, 16(3), 4.

Wise, D. A. Prediction of test scores: California. Washington, D. C.: U. S. Department of Health, Education, and Welfare, 1976. (ERIC Documents Reproduction Service No. ED160306)

Wolfley, L. M. Strategies of path analysis. American Educational Research Journal, 1980, 17(2), 183-209.

Wolfley, L. M. A method for estimating indirect effects in path analysis. Paper presented at the annual meeting of the American Education Research Association, New York, March 19, 1982.

Worth. W. H. The effect of promotion and nonpromotion on pupil achievement and social-personal development in the elementary school (Doctoral dissertation, University of Illinois, 1959). Dissertation Abstracts International, 1960, 20, 3228. (University Microfilms No. 60-256)

Worthington, L. H., & Grant, C. W. Factors of academic success: A multivariate analysis. Journal of Education Research, 1971, 65, 7-10.

Yee, A. H., & Gage, N. L. Techniques for estimating the source and direction of causal inference in panel data. Psychological Bulletin, 1968, 70, 115-126.

Appendix A

Number of Schools Organized with Grades 1 and 3 on the Same Campus
Mean and Standard Deviation for N.C. School Systems
Data From 1980-1981 North Carolina Education Directory

Number of LEA's	Number of Schools/LEA	$(x_i - \bar{x})$	$(x_i - \bar{x}.)^2$	$n(x_i - \bar{x}.)^2$
5	0	-7.586	57.549	287.736
13	1	-6.586	43.375	563.880
13	2	-5.586	31.203	405.644
10	3	4.586	21.031	210.314
14	4	3.586	12.859	180.032
12	5	2.586	6.687	80.249
10	6	1.586	2.515	25.154
12	7	.586	.343	4.121
13	8	.414	.171	2.228
4	9	1.414	1.999	7.998
8	10	2.414	5.827	46.619
7	11	3.414	11.655	81.588
3	12	4.414	19.483	58.450
3	13	5.414	29.311	87.934
4	14	6.414	41.139	164.558
4	15	7.414	54.967	219.870
4	17	9.414	88.623	354.494
1	21	13.414	179.935	179.935
1	27	19.414	376.903	379.903
1	29	21.414	458.559	458.559
1	33	25.414	645.871	645.871
1	47	39.414	1553.463	1553.463
1	58	50.414	2541.571	2541.571
145	1100			8540.171/145=
				58.898

S.D. = 7.67
Mean = 7.586
Median = 6
Mode = 4

Appendix B

	Ach X_1	Ret X_2	IQ X_4	Race X_5
Ret X_2	r=0 (2) r=0 (17)			
IQ X_4		r= - of unknown value of (16, 17)		
Race X_5	r=.34 (9)	r=0 (18) r=0 (16)	r=.90 (13)	
SES X_6	r=.50 (13) r=.70 (9)	r=0 (18)	r=.86 (13)	r=-.2, -.6 (10) r=-.58, -.79 (13)

Figure 6. Expected Correlation Matrix of Model Variables from Aggregate Data

	Ach X_1	IQ X_4
Ret X_2	r=0 (1)	
IQ X_4	r=.57, .67 (3) r=.79 (4) r=.68, .84 (5) r=.46, .82 (6) r=.78, .85 (7) r=.61, .65 (8)	
Race X_5		r=.36 (20)
SES X_6	r=.44 (12) r=.34 (5)	r=.30, .32 (5) r=.68 (6) r=.33, .49 (12)

Figure 7. Expected Correlation of Matrix of Model Variables from Individual Data

Index of Authors for Expected Correlation Matrices

1. Jackson
2. Coffield
3. Asbury
4. Arena
5. Knief
6. Baker
7. Hawkes
8. Brookover
9. Innes
10. Wise
11. Fort
12. Deutsch
13. Burkhead
14. Conlisk
15. Lloyd
16. Clark
17. Cook
18. Craig
19. Rogers
20. Backman
21. Semler
22. Scarr-Salapatek

FARMER SCHOOL

FARMER RURAL STATION

Asheboro, N. C. 27203

Appendix C

SAMPLE REQUEST FOR DATA

I am in the process of completing a doctoral dissertation at the University of North Carolina at Greensboro. I propose to investigate the relationship between retention rate and achievement in elementary schools. Most previous research in this area has examined the effect that retention-in-grade has upon the retained child; little attention has been paid to the effect of retention on the majority of children in the class who were not retained.

One set of educational theorists would argue that failing damages a child's self-image and decreases his willingness to attempt academic work. Therefore, schools with higher rates of retention should, in general, demonstrate lower mean achievement. Conversely, other educators argue that strict establishment of performance standards and retention of pupils who do not perform will improve achievement by giving students a goal to strive toward.

Unlike previous research, this paper will examine the relationships between mean achievement, per grade, per school and the percentage of pupils retained in the previous grades.

The State Department of Public Instruction (SDPI), Division of Annual Testing, has reviewed my dissertation proposal and has agreed to provide mean achievement data for North Carolina Schools, grades 1, 2, and 3 from the spring test results of 1979, 1980, and 1981. They will also identify the socio-economic and racial compositions for each group. Retention rates per school, per grade are part of each principal's annual report. As you can see, virtually all of the data for this investigation are available from and are being provided by SDPI. However, to properly define the relationships among these factors, I must also consider school-by-school ability data.

Because your school system is organized to facilitate analysis, I would like to include it in the study. I discussed this proposal with you on June 20, 1981, and was encouraged by your cooperation. The data which I need are the third grade mean I.Q.'s for 1980-81 for each school in your system which serves grades one through three. For your convenience I have enclosed a data sheet with school names and numbers already filled out. Although I must identify schools in my data gathering to assure that I properly synthesize information from you and the State Department, at no time in the reporting of my results will I identify the participating schools.

If you should have any questions concerning this request, please feel free to call me at Farmer School (919-857-2156) or to contact Dr. Richard Jaeger at UNC-Greensboro (919-379-5517) who is serving on my committee. Thank you in advance for your prompt attention to this request.

Yours truly,

John C. Maddocks

Enclosure