DOCUMENT RESUME

ED 271 510 TM 860 421

AUTHOR Kerckhoff, Alan C.
TITLE Effects of Ability Grouping in Secondary Schools in Great Britain.
INSTITUTION Ohio State Univ., Columbus. National Center for Research in Vocational Education.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE 85
GRANT NIE-G-83-0005, P-1
NOTE 50p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Ability Grouping; *Academic Achievement; Analysis of Variance; Cohort Analysis; Educational Background; Elementary Secondary Education; Foreign Countries; Longitudinal Studies; *Mathematics Tests; Parent Influence; *Reading Comprehension; Regression (Statistics); *School Organization; *Scores; Sex Differences; Social Background; Tables (Data)
IDENTIFIERS *Great Britain; National Child Development Study

ABSTRACT

This paper reports the findings of an investigation of the effects of school organization and ability grouping on students' academic achievements in Great Britain. Data for the study came from the National Child Development Study (NCDS) conducted by the National Children's Bureau (NCB) of London. The NCDS surveyed virtually every child born in England, Scotland, and Wales during the week of March 3-9, 1958, and the present study included data from follow-up studies conducted when the cohort was 7, 11, and 16 years old. Types of schools attended by the cohort were: (1) comprehensive; (2) grammar; (3) secondary modern schools in the state sector, and (4) private schools. Some, but not all, of the schools practiced ability grouping. Measures included as control variables were social background, parent influences, and school influences. Results of the study showed that separation of students into ability groups had an effect on achievement test performance in both reading and mathematics. The four types of schools received students whose average earlier test performance varied systematically in ways that were consistent with the social definitions of the schools. The several types of ability groups included students whose earlier performance suggested they had the designated levels of ability. (LMO)

EFFECTS OF ABILITY GROUPING IN SECONDARY SCHOOLS IN GREAT BRITAIN

by

Alan C. Kerckhoff
Professor of Sociology
and
Chair of the Department
Duke University

The National Center for Research in Vocational Education
The Ohio State University
1960 Kenny Road
Columbus, OH 43210

1985

FUNDING INFORMATION

Project Title:	Schooling Effectiveness
Grant Number:	NIE-G-83-0005, P-1
Project Number:	716864
Act Under Which Funds Were Administered:	P.L. 96-88
Source of Contract:	U.S. Department of Education National Institute of Education Washington, DC 20202
Project Officer:	Ronald Bucknam
Contractor:	The National Center for Research in Vocational Education The Ohio State University Columbus, OH 43210-1090
Executive Director:	Robert E. Taylor
Disclaimer:	This publication was prepared pursuant to a contract with the National Institute of Education, U.S. Department of Education. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official U.S. Department of Education position or policy.
Discrimination Prohibited:	Title VI of the Civil Rights Act of 1964 states: "No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving federal financial assistance." Title IX of the Education Amendments of 1972 states: "No person in the United States shall, on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving Federal financial assistance." Therefore, the Research Program on Employability, like every program or activity receiving financial assistance from the U.S. Department of Education, must be operated in compliance with these laws.

FOREWORD

This paper is a product of the nonresident scholar program conducted by the Research Division of the National Center for Research in Vocational Education. This program is designed to draw on scholars from outside the National Center in order to assure a broad coverage of research topics and spark a lively exchange of ideas among top scholars in the country.

Thanks are due to the author of this paper, Alan C. Kerckhoff, Professor of Sociology and Chair of the Department, Duke University. Additionally, appreciation is expressed to Lawrence Hotchkiss, Research Specialist, the National Center for Research in Vocational Education and to John Bishop, Associate Director of the Research Division at the National Center for coordination of the nonresident scholar program.

EFFECTS OF ABILITY GROUPING IN SECONDARY SCHOOLS IN GREAT BRITAIN

by

Alan C. Kerckhoff
Duke University

Numerous investigations have been conducted over the past two decades to determine the effects of the characteristics of schools attended on the academic performance of students. Since the "Coleman Report" (Coleman, et al., 1966) produced negative findings, little interest has been shown in such school characteristics as the nature of the physical plant or even the kinds of educational facilities (e.g., laboratories, libraries, etc). The more recent literature has been more concerned with the characteristics of the student body or the internal organization of the school. Some evidence has been presented to support the belief that the general characteristics of the student body affect the academic performance of individuals (Alexander and Eckland, 1975). But the claim that the school context does not affect all students in the same way (Heyns, 1974) has also directed attention to the internal organization of the school. While the internal organization of a school is part of the overall school context experienced by students, a student's experience is quite different depending on where s/he is located in that structure.

This paper reports the findings of an investigation of the effects of school organization on students' academic achievement in Great Britain. It is concerned with the effects of what most Americans call "tracking" and the British refer to as "streaming," the segregation of students in groups defined by ability levels. While the approach taken is consistent with that used in studies of American schools, differences between the British and American school systems and some methodological innovations make it possible to increase our understanding beyond what is possible from previous research reports.

Early attempts to determine if the segregation of students into ability groups leads to performance differences which would not be observed without the segregation have produced mixed findings and differences of opinion about the proper interpretation of the findings (Esposito, 1973; Findlay and Bryan, 1975). There have been those who saw ability grouping as a mechanism by which schools artificially directed students toward varying levels of adult outcomes (Rosenbaum, 1976). On the other hand, not everyone thought that the outcomes which could be associated with ability grouping were sufficiently evident to warrant any concern (Jencks, et al., 1972).

Those who have believed that ability grouping does affect performance have thought that the effect is to increase the level of performance of those in the high ability groups and to lower the level of performance of those in the low ability groups in comparison to what it would have been without the separation into ability groups. The explanations suggested for expecting such effects varied. Rosenbaum (1976) suggested that those in the lower tracks were discriminated against in the grading practices used and in the way school records are kept as well as being given an inadequate curriculum. Oakes (1982) also stresses the importance of the poor curriculum in the lower tracks as well as the teachers' lower expectations of the students. Alexander, et al. (1978) add to this the fact that ability level segregation places students in varying contexts differentiated by the kinds of peers with whom they associate; if such contextual effects are significant at the school level, they should be even more significant at the learning group level.

Two specific methodological critiques have been aimed at the research on this problem. One of these is essentially a "friendly" criticism. It suggests that since some of the research in this area has used students' self reports of track placement, the validity of the classification is questionable. It notes

that there is less than perfect agreement between such student reports and official records of track placement. Most important, it has been shown (Rosenbaum, 1980) that track defined by the official records has a stronger effect on some outcomes, such as college attendance, than does self-reported track. In effect, this criticism suggests that there is considerable measurement error in the self-reports, which weakens any possible effect they could exhibit in the analyses conducted.

The second criticism is even more basic and clearly important. Early studies of the effects of ability grouping were not always able to control for the characteristics of the individuals before they were put into ability groups. It could thus be argued that what purported to be effects of grouping (Alexander and McDill, 1976) were actually effects of the students' characteristics prior to entering academic tracks. In effect, this criticism suggests that the analytic models used in the early research were misspecified. Some more recent analyses, controlling for pre-tracking performance (Alexander, et al., 1978), have shown that this improved specification does reduce the effects of ability grouping, but they also demonstrate that significant ability grouping effects remain. While not all recent research shows such residual effects (Rehberg and Rosenthal, 1978), the evidence seems to be mounting.

Probably the most intensive study of this issue, using the large sample and multiple measures of the High School and Beyond data set (Hotchkiss, 1984, Chapter 7), has recently provided additional support for the hypothesis that placement in tracks does have an effect on academic performance, net of a long list of control variables measured at an earlier point. That analysis has the limitation of covering only the period from tenth grade to twelfth grade, but its findings are very impressive and wholly in keeping with the earlier research which has indicated significant effects of ability grouping. In a

later section of this paper, reference will be made to Hotchkiss' findings in comparison with those to be presented below.

Before turning to the results of the present study, however, it is worthwhile to indicate a limitation of all of the recent studies of ability grouping effects. The studies cited above use a multiple regression format for the analysis. Following the convention of assessing change across time by measuring the lagged effects of earlier measures of the dependent variable, they have determined whether ability group assignment contributes to the explanation of the dependent variable, net of the lagged effects (and any control variables included). This kind of approach is certainly appropriate and will be used in the present analysis. However, previous investigations have ignored the distributional characteristics of the several kinds of ability groups. Although the regression analysis can determine whether or not group placement tends to increase or decrease performance, compared with what would otherwise be expected, it does not consider the actual levels of performance at either the earlier or later points of measurement, nor does it consider the degree of variation within groups at various levels.

The theoretical stance implicit in most, if not all, studies of the effects of ability grouping provides a basis for making predictions about both mean levels and distributions of performance measures in various kinds of ability groups. To make that theoretical position clear, it may help to refer to the work of Wohlwill (1973) and Boudon (1974) who have presented very similar conceptualizations of the over-time effects of social settings, although their substantive interests were rather different. Wohlwill makes reference to a "positive feedback model" of situational effects while Boudon refers to an "exponential effect" of persisting social contexts. In both cases, they refer to the expectation that if two individuals are consistently exposed to different social contexts, and if those contexts can be expected to

affect an individual's personal characteristics in opposite ways (e.g., one will increase academic performance, the other decrease it), over time we should expect a progressively increasing difference in the two individuals' characteristics.

If ability groups actually have the hypothesized effects (i.e., high ability groups increase performance, low ability groups decrease performance), we would thus expect greater mean differences between groups at the end of their segregation than at the beginning. If that outcome is actually observed, two other expectations might be derived from it, although they are to some extent contradictory. First, if the performance levels of the ability groups diverge over time, it is logical to expect the variance in performance of individuals included in the larger unit (e.g., the school) of which the segregated groups are parts to be greater at the end of the period of segregation than at the beginning. However, if being in an ability group has a dominant effect on all the group's members, continued exposure to the group setting could lead the group members to become more homogenous by the end of the period of segregation than they were at the beginning.

In the present analysis, therefore, it will be important to examine the means and variations of performance of the groups of students both prior to and after they have been separated into ability groups. Diverging mean performance levels between high and low ability groups would be expected. In addition, one or both of two changes in variation of performance would be expected. Variation of performance levels within ability groups may decline, and/or variation within types of schools which streams may increase.

Sample and Method

The present analysis of ability grouping in Great Britain has two major advantages over previous studies. First, it is based on a birth cohort whose

lives have been charted from birth to early maturity and for whom there is thus information from earlier ages than is usually available for such analyses. Second, the British school system at the time these individuals were in secondary school had a greater variety of organizational arrangements than have been considered in previous studies of ability grouping. In order to put the subsequent analysis in perspective, it will be necessary to present a few introductory statements about the sample and the school system.

The data come from the National Child Development Study (NCDS) conducted by the National Children's Bureau (NCB) of London. The NCDS originated as the "Perinatal Mortality Survey," a study of "virtually every baby born in England, Scotland and Wales during the week of 3 to 9 March [1958]" (Davie, Butler and Goldstein, 1972, p. 10). Follow-up studies were conducted when the cohort was 7, 11, 16 and 23 years old. The present analysis includes data from the 7, 11 and 16 year old time points, but the majority of the measures were made when the cohort was 11 and 16 years old.

The core measures are achievement tests in reading and mathematics specially administered in the schools for the NCDS when the cohort was 7, 11 and 16 years old. At age 7, the Southgate Reading Test and a Problem Arithmetic Test (specially designed by the National Foundation for Educational Research [NFER]) were administered. At age 11, a Reading Comprehension Test and an Arithmetic/Mathematics Test (also specially designed by NFER) were administered. At age 16, the same Reading Comprehension Test was administered as at age 11, and a Mathematics Test devised at the University of Manchester was administered. Also, at age 11, a general ability test was administered from which verbal and non-verbal scores were derived.

The age 11 tests were administered during the last year of junior school, just prior to the students' transfer to secondary school. (This is actually a

year prior to transfer to secondary school in Scotland where the change of schools occurs at age 12 rather than 11.) The measures may thus be considered indications of the students' ability and achievement levels prior to experiencing ability grouping in secondary school. The age 16 measures were administered just prior to the earliest point at which the students could leave school and thus the last point at which the full cohort could be administered tests. (Nearly two-thirds of the cohort left school at age 16.)

The vast majority of the cohort attended one of four kinds of secondary schools. After World War II, there were three kinds of British state-supported secondary schools, the grammar, technical and secondary modern schools. Grammar schools were attended by students of high ability who were deemed suitable to prepare for university attendance. Students who were gifted but who were destined for technical occupations were likely to attend technical schools. Other students, not chosen for either grammar or technical school attendance, attended secondary modern schools. By 1969 or 1970, when the cohort moved to secondary schools, a new form of school, the comprehensive school, had become predominant, and over half of the cohort attended one. It was not a selective school, but was intended to enroll a cross-section of students. Also, by that time, the technical school had almost disappeared. But both grammar and secondary modern schools still enrolled significant proportions of the students. Thus, the four types of schools attended by the cohort were comprehensive, grammar and secondary modern schools in the state sector and private schools. The analysis includes all students who attended those four types of schools and for whom full data are available.

Some, but not all, of the schools of all four types practiced ability grouping. The British differentiate between two kinds of grouping, "streaming" and "setting." Streaming refers to the separation of students by ability level

for all of their classes while setting refers to separation for only particular classes, such as science or mathematics. In the analysis presented here, no distinction is made between the two. Instead, all students who were separated into ability groups of either kind for mathematics or English were classified together.

The schools were asked whether each student was in an English or mathematics class grouped by ability and, if so, which group s/he was in. The groups could be classified into six types. In some schools, there were three levels of ability groups and in others only two. It was thus possible for a student to be in a high or low group in either a two- or a three-group system or in a middle group in a three-group system. In addition, some students were in what were referred to as remedial groups. Therefore, in addition to students who were in ungrouped classes, there were those in remedial classes and those in high, medium and low levels in a three-group arrangement and those in high and low levels in a two-group arrangement.

The analysis to be conducted is thus much more refined than has been carried out in the United States. It differentiates among four kinds of secondary schools and, within each of those, among seven kinds of groupings of students rather than a simple comparison between two groups (usually college preparatory versus all others) as in almost all American studies. In all of the analyses, the effects of ability grouping can be assessed for the several levels and kinds of grouping in comparison with students in the same kind of school who did not experience ability grouping.

In addition to these core test measures and the school and ability grouping categories, other measures are included in the analysis as control variables. They are the following:

Social Background

Social class of father's occupation (seven categories) when the child was 7 years old.

Social class of father's occupation when the child was 11 years old.

Age at which father left school.

Age at which mother left school.

Whether the child was nonwhite.

Whether the child lived with his or her natural or adoptive mother.

Whether the child lived with his or her natural or adoptive father.

Parent Influences

Whether, when the child was 11, the parents wished the child to leave school as soon as possible.

Whether, when the child was 11, the parents wished the child to seek some kind of education beyond secondary school.

School Influences

Whether the child attended a private school at age 7.

Whether the child attended a private school at age 11.

The teacher's rating of the child's reading ability at age 7.

The teacher's rating of the child's mathematics ability at age 7.

The teacher's rating of the child's use of books at age 11.

The teacher's rating of the child's mathematics ability at age 11.

Contingencies

Number of schools attended between age 5 and age 11.

Number of residential moves between birth and age 11.

Child's school attendance at age 7.

Child's school attendance record during 1972 (when the cohort was 14 years old).

Whether the child has some kind of handicap.

A total of 9,399 individuals (4,797 boys and 4,602 girls) are included in the analysis. While there have been losses from the sample over the years, as happens in any longitudinal study, the losses have not had a strong biasing effect on the sample characteristics. A general indication of the amount of bias due to missing data is provided by Goldstein (1976, p.70) based on a detailed analysis of the data through age 16:

> Children who have previously belonged to 'disadvantaged' groups are less likely to provide any information at 16 years, and for estimates of the proportions of children with particular characteristics an upper limit for the relative bias at 16 years is about 10 percent. . . . For mental test scores at 11 years, the differences between the extreme categories of variables such as social class and family-size ratings have small biases of up to about 3 percent. . . . There is evidence in the case of mathematics attainment of a higher rate of change in attainment between 7 and 11 years for those with data as 16 compared with those without. A simple extrapolation of this difference to 16 years gives an overall bias in mathematics attainment of about 0.05 years, which would seem to be acceptably small.

Findings

The results of the analyses will be presented in two parts. The first part will be concerned with the levels of performance of individuals in the various groups, at both age 11 and age 16. The second part will present the results of the regression analyses.

Levels of Performance

Means and standard deviations of reading and mathematics test scores at age 11 and 16 were computed for boys and girls separately for each group formed by the combination of school type and ability grouping arrangements. The means and standard deviations of the full cohort, the four school types and the seven ability groups are presented in Tables 1 and 2. The full set of tables of these statistics, showing the scores of ability groups within school types, can be found in Tables A1 through A4 of the Appendix.

For present purposes, several general observations can be made about these data. First, the mean levels of performance in both reading and mathematics at both ages are consistent with expectations, given the divisions of the cohort into school types and ability groups within schools. Average scores of students in comprehensive schools are most similar to the average for the full cohort - although they are slightly lower, suggesting that the top level students are "creamed" (to use a British term) by the grammar and private schools. Students in both of the latter types of schools score well above, and those in secondary modern schools score below, those in comprehensives (see Table 1).

Tables 1 and 2 about here

Similarly, the sub-group average scores follow the expected pattern. Those who were not in ability groups have average scores very similar to the full sample. (The exception is the ungrouped girls' average mathematics score, which is clearly lower than the overall girls' average mathematics score.) Remedial students score the lowest of any sub-group, at both ages, and the means for the five types of ability groups fall in the sequence expected - even to the extent that the high and low groups in a three-group system having more extreme average

scores than the high and low groups in a two-group system. In the great majority of cases, these patterns are found within school types as well as for the full sample (see Appendix Tables A1 - A4.)

The earlier discussion presented a major hypothesis and two possible derivative hypotheses relevant to these data. The core hypothesis of the research is that average scores of students in high ability groups should increase and those of students in low ability groups should decrease during the secondary school years, relative to what they would have been had the students not been segregated into ability groups. The regression analysis to be presented below is the most convincing way to test that hypothesis because it can take earlier levels of performance and other variables into account. However, a first impression of that matter may be gained from these data.

Only in the case of the reading test can a simple comparison be made between the test scores at the two points since the same test was given both times. However, since the purpose of the present examination of the data is to determine if there are systematic differences across groups in the amount of change, it is helpful to make such comparisons for both reading and mathematics scores. The word "change" will be used for convenience when discussing both reading and mathematics, but it should be remembered that the differences in mathematics scores at 11 and 16 do not actually refer to changes on a single scale the way the differences in reading scores do.

Table 3 reports the changes in means of the reading and mathematics test scores between age 11 and age 16 for school types and ability groups, based on the information in Tables 1 and 2. If all types of schools and all ability groups produced the same results, we would expect all of the mean changes to be approximately the same as the changes reported in the "Total" row of the table. The number entered for each school and ability group is the difference between that group's mean change and the total sample mean change.

Table 3 about here

Looking first at the reading scores, it is apparent that there are differences in the average changes across school types and ability groups, but not all of them are what we might expect. Students in comprehensive schools and students who were not in ability groups come closest to reflecting the total changes for both sexes. The differences among school types are not great, but, for both sexes, those in secondary modern schools gained more than any of the others, the reverse of what one might have expected, given the general hypothesis guiding the analysis. Yet, the differences between school types are so small, they could easily be accounted for by a tendency for scores to regress toward the overall mean. School types with higher age 11 scores gain less than those with lower age 11 scores.

The reading score differences among ability groups are nearer to those called for by the hypothesis. The gains in the high and low ability groups in both the 3- and 2-group structures fit the expectation. They are, respectively, above and below the total sample gains and the gains for ungrouped students. The remedial group students gain less than any other group. Since these differences are the reverse of what would be expected if there were regression toward the overall mean, they suggest some support for the hypothesis.

The results for mathematics scores are rather different from those for reading scores. The differences among school types are generally the same as for reading scores - students in secondary modern schools deviate in a positive direction and those in grammar and private schools deviate in a negative direction from the overall sample. While these differences are also consistent

with a regression toward the mean interpretation, they are more sizeable than those for reading

The mathematics mean differences for ability groups also fit a regression toward the mean pattern. It is clear that the remedial groups have the most positive deviation, and the high ability groups have the most negative deviation. Again, these differences are rather large.

Not only are the results different for reading than for mathematics, they are almost reversed. High ability reading groups gain more than low ability groups, while high ability mathematics groups seem to gain less than low ability mathematics groups. Further comment on this difference will be postponed until after the results of the regression analysis have been presented.

One of the derived hypotheses called for an increase in the variance of test scores within school types between age 11 and age 16, due to the expected tendency for high and low ability group scores to diverge. Again, this hypothesis can most easily be assessed by examining the patterns of reading scores, but in order to obtain a basis for some comparison between the effects for mathematics and reading scores, both will be considered. Since different mathematics tests were administered at 11 and 16, and since they led to very different score means and standard deviations, the analysis of the mathematics data must be viewed as providing only a crude indication of changes over time.

The top panel of Table 4 reports the differences between the standard deviations of the reading and mathematics scores of the total sample and the four school types. Again, the raw difference between the age 11 and age 16 standard deviations is entered for the full sample, and the deviation from that full sample difference is shown for each school type. Table 4 indicates that there was a small increase in the standard deviations of reading scores of the full samples of both boys and girls. Roughly parallel gains are seen in comprehensive schools. What is most striking, however, are the very different kinds

of change in reading standard deviations in the other school types than in the comprehensive schools. The secondary modern school students experience a much greater increase in standard deviation while both grammar and private school students show a sharp reduction.

As before, we find the pattern is rather different for mathematics. Although the standard deviation changes for comprehensive school students are again most similar to those for the full sample, the other school types deviate from the comprehensive school change in different ways than they did for reading scores. While standard deviation changes for students in secondary modern schools are very similar to those for comprehensive school students, the standard deviations of grammar school students are reduced less than those of students in any other type of school. This is in sharp contrast to the reading scores, where the grammar school standard deviations were reduced more than for any other type.

The result of these differences is that, across school types, the standard deviations of mathematics scores are more similar in size at age 16 than they were at 11, while the standard deviations for reading scores are less similar in size at 16 than they were at 11. For neither reading nor mathematics, though, is there consistent support for the derived hypothesis calling for an increase in the sizes of standard deviations within school types. There is an increase in reading standard deviations in comprehensive and secondary modern schools, but there is a clear decrease in grammar and private schools. It may also be possible to claim that there is an increase in the mathematics standard deviation in grammar schools, but it is certainly not found across the four school types.

The second derived hypothesis calls for a decrease in standard deviations within ability groups. The bottom panel of Table 4 presents the relevant data

in a form parallel to the top panel. Here, as with comprehensive school students, we find the changes in reading standard deviations of the ungrouped students to be similar to the totals reported at the top of the table. But again we find that students in different kinds of ability groups exhibit very different changes in standard deviations between the two testing times. Students in high ability reading groups become more homogeneous at 16 than they were at 11, while students in low ability groups (and especially the remedial groups) become more heterogeneous.

Table 4 about here

The mathematics data again provide a very different picture. The standard deviation differences for remedial groups suggest that they become more heterogeneous, which is parallel to the finding for reading scores. But the differences for the other ability groups suggest the opposite pattern than found for reading scores. High ability groups appear to become more heterogeneous and low ability groups less hetereogeneous between 11 and 16. Since the mathematics measures are not the same at the two time points, these differences are only suggestive, but they do not appear to follow the same pattern as the reading scores.

There is thus no consistent support for the second derived hypothesis calling for increased homogeniety within ability groups. Even if only the changes in reading score standard deviations are considered, it is apparent that the second derived hypothesis is not consistently supported. There is no general tendency for the variation in scores within ability groups to decrease over the five-year period. So far as reading is concerned, this seems to be the case for high ability groups but not for low ability groups. While the mathematics data are not wholly adequate for this purpose, the opposite seems to be

true - variation in high ability groups seems to increase more than that in low ability groups.

Given the limited nature of the available data, it is not possible to determine the reasons for these different effects across school types, ability group levels and subject matter. However, in the final section of this paper I will suggest an interpretation of these findings which seems at least worthy of consideration and further investigation. The analysis thus far has not adequately dealt with the major hypothesis of the research. Further discussion of the patterns of means and standard deviations will be postponed until the analysis relevant to that hypothesis has been presented.

Modeling Changes in Achievement

The core hypothesis being considered in this paper is that ability grouping tends to increase the level of performance of students in high ability groups and decrease the level of performance of students in low ability groups compared with what they would have been if the students had not been segregated into ability groups. While some of the results presented above are at least consistent with that hypothesis, they do not in any effective way provide a test of it.

In order to deal with that hypothesis more effectively, we regressed the age 16 test scores on three prior test scores. The age 16 reading scores were regressed on the age 11 and 7 reading scores and the age 11 general ability (verbal) test score. The age 16 mathematics score were regressed on the age 11 and 7 mathematics scores and the age 11 general ability (non-verbal) score. The three prior scores thus represent the best basis we have for predicting the age 16 scores.

In each regression analysis, dummy variables were used to represent either the school type or the ability group of the individual students. In analyses in

which school types were included, the comprehensive school was used as the reference category. Thus, the coefficients for the other school types represent the average difference in age 16 achievement due to having attended that type of school rather than a comprehensive school. Similarly, in analyses in which ability group dummy variables were used, the ungrouped students were used as the reference category. In those analyses, therefore, the coefficients of the ability groups represent the average difference in age 16 achievement due to having been in a particular kind of ability group rather than having been in an ungrouped class.

As with the analysis of means and standard deviations, the regressions were computed for the full set of possibilities - using only school types, using only ability groups, and using ability groups within school types. The full, detailed tables for ability groups within school types are in Appendix Table A5. Table 5 presents the results from the analysis of the full sample. The metric coefficients for school types and ability groups are reported as well as the adjusted R^2 of the full equation and the adjusted R^2 for the equation using only the earlier test scores as the independent variables.

Table 5 about here

The analysis of school types shows a sharp contrast between the results for reading and for mathematics. School type has essentially no effect on age 16 reading scores, net the characteristics of the students prior to entry into secondary school. On the other hand, school type has a very significant effect on age 16 mathematics test scores. Having attended a grammar school or a private school adds between two and three points to the average score on the age 16 mathematics test beyond what would be expected, given the students' earlier test performance. That represents between one-third and and one-half a standard

deviation of the test scores (see Table 1) - a very considerable gain. That is an especially impressive gain, given the fact that the grammar and private school students' scores were the highest overall when they entered secondary school.

The effects of ability grouping are apparent in both reading and mathematics. The effects on reading performance are most visible for the remedial reading group. In spite of their very low average performance at age 11 (see Table 2), they lose a great deal of additional ground during secondary school. It is not at all clear that any remediation has occurred. (In addition to the average loss indicated in Table 5, Table 2 also shows that the variation in scores has increased by age 16, suggesting that some of the remedial readers have fallen very far behind their age mates.) The remedial group has come closer to holding its own in mathematics, although the boys in the group have slipped back to a significant degree.

For those in the other five ability groups, there are generally significant effects of the sort called for by the core hypothesis being considered here. Those in the high ability groups score significantly higher than expected, given their earlier test performance, and those in the low ability groups score significantly lower, in comparison with students who were not in ability groups. The spread created by gains in high and losses in low ability groups is greater for girls in reading and for boys in mathematics. Also, for both sexes and in both test areas, the spread is greater for three-group than for two-group systems. The spread in the three-group systems is over four points for girls' reading scores and for boys' mathematics scores. These are about two-thirds of a standard deviation of the age 16 test scores (see Table 1), a very large effect.

Before accepting these results as definitive, it is important to raise the further question of other possible sources of the observed changes over the five-year period. As indicated earlier, four sets of variables, all antecedent to the age 16 testing, were considered. They were measures of the family background of the student, the student's parent's wishes for his or her educational attainment, additional measures of earlier school influences, and several contingent factors that might be expected to affect academic performance. These four sets of variables were included individually and in combination in the regression analyses to control for possible extraneous influences, not directly relevant to the effects of ability grouping.

As before, these analyses were conducted for the full sample with school type represented by dummy variables, with ability groups represented by dummy variables, and within school types with ability groups represented by dummy variables. The full results, including those within school types, are presented in Appendix Table A6.

Table 6 presents the coefficients for school types and ability groups from the analysis using all of the control variables in a form parallel to Table 5. A comparison between the two sets of coefficients indicates that, although the control variables have added somewhat to the explanation of variance in age 16 test scores, and although the sizes of the coefficients for school types and ability groups have been reduced somewhat, the form of the results is essentially the same as without the controls.

Table 6 about here

School type continues to have little effect on the level of performance in reading, although there is an interesting reversal in the signs of the boys' coefficients for grammar and secondary modern schools. Boys who attend grammar

schools fall below their expected level of performance, and those who attend secondary modern schools rise above their expected level. The direction of these effects was the same in Table 5, but the coefficients were not statistically significant until the control variables were introduced. I have no ready explanation for this result, but since it is the reverse of the finding for mathematics, it may suggest that the two types of schools emphasize different subjectmatter, at least for boys.

Those who attend grammar or private schools gain significantly in mathematics performance beyond what would be expected from their earlier test scores and all of the other antecedent conditions considered. While the size of the gain has been reduced over that reported in Table 5, it still represents a gain of between one-fifth and one-third of a standard deviation in the age 16 mathematics test scores.

The results for ability groups continue to be even more significant than for school types. The remedial reading group is still the most disadvantaged of all, falling from three-fifths to nearly three-fourths of a standard deviation below their expected performance, given the full set of antecedent variables. There is no disadvantage for the remedial groups in mathematics, however, the girls actually performing somewhat better than expected. Among the other ability groups, there are still sizeable differences between the coefficients for the high and low groups. As before, the high-low spread is greater for the three-group comparison, and the spread is greater for girls in reading and for boys in mathematics. The three-group spread in reading for girls and in mathematics for boys are both in excess of one-half standard deviation.

Although these results for ability grouping are highly significant and have withstood the introduction of a large number of additional possible sources of influence, they are not fully consistent across school types. (The full set of

coefficients is presented in the Appendix Tables A5 and A6.) Several variations across school types are noteworthy:

(1) There are very few individuals in remedial groups in either grammar or private schools. Thus, the strong effect of remedial group segregation applies only to comprehensive and secondary modern schools, and it tends to be stronger in the former than the latter.

(2) Very few of the individual ability group coefficients are statistically significant in the grammar or private school type analyses. Also, although most of coefficients follow the expected pattern, there are a number of reversals from the general pattern in the order of the coefficient sizes and signs. These are more common for mathematics in grammar schools and for reading in private schools.

(3) Overall, therefore, the expected pattern of increased scores for high ability groups and decreased scores for low ability groups is found most clearly and consistently in comprehensive and secondary modern schools.

Discussion

The evidence presented leaves little doubt that separation of students into ability groups has an effect on performance on achievement tests in both reading and mathematics. The four types of schools receive students whose average earlier test performance varies systematically in ways that are consistent with the social definitions of the schools (Table 1), and the several types of ability groups do include students whose earlier performance suggests they have the designated levels of ability (Table 2). Even within the selective schools,

the ability groups effectively separate students according to levels of earlier performance (see Appendix Tables A1 through A4).

The regression analyses graphically demonstrate the effects of ability grouping, even when a great many other sources of possible influence are controlled. Students in remedial classes lose a great deal of ground (at least in reading). Students in low ability groups lose ground, and those in high ability groups increase their average performance level beyond that exhibited by students in ungrouped school settings. The losses by low ability students, combined with the gains by high ability students, make the overall effect of ability grouping very striking. The pattern is so clear that it is even possible to differentiate between the effects of a two-group and a three-group system, the latter producing a greater high-low ability group spread.

The major hypothesis investigated in this analysis has thus received considerable support. While there is evidence of the effects of ability grouping in all four school types, it is much clearer in comprehensive and secondary modern schools than in either grammar or private schools (see Appendix Tables A5 and A6). At the same time, it is noteworthy that there is a significant overall gain in mathematics performance associated with attending a grammar or private school (see Table 6). As with the ability groups differences, this is especially impressive given the fact that students entering those types of schools have the highest average achievement scores at the outset.

The contrast between grammar and secondary modern student outcomes in mathematics parallel those between high and low ability groups. Since students are chosen to attend these two types of state supported schools on much the same basis as students are chosen for high and low ability groups within school, this contrast also tends to support the major hypothesis of the research.

The analysis did not provide any consistent support for either of the derived hypotheses. There is not a general tendency for the standard deviations to increase in size within school types nor to decrease in size within ability groups.

Two other variations were observed in the findings. There is a sex difference in the effects of ability grouping. The effects of grouping in reading are greater for girls, and the effects of grouping in mathematics are greater for boys. This difference by subjectmatter is consistent with other sources which suggest girls' superiority in reading and boys' superiority in mathematics, but it is also consistent with the hypothesis that during adolescence there is a general tendency for girls' mathematics performance to fall behind that of boys. This tendency has been noted in both the U.S. and Great Britain (Douglas, et al., 1968; Grant and Eiden, 1982). While the present results are consistent with that observation, they do not directly provide substantial support for it.

Even more impressive than the sex differences have been the consistent differences in the results between the reading and mathematics analyses. These differences have been noted in the discussion of mean changes and changes in the standard deviations of the measures as well as in the discussion of the regression analyses. These discussions have repeatedly noted that the results differed for the two subjects across ability groups and school types. There seems to be some consistency in these differences, and it may be worthwhile to discuss it here.

In the analysis of reading scores, there is a clear tendency for grammar and private school standard deviations to decrease and those for secondary modern schools to increase between 11 and 16. A parallel pattern is found for ability groups, high ability groups becoming more homogeneous while low ability groups become more heterogeneous. In mathematics the opposite seems to occur; the

standard deviations for private and (especially) grammar schools increase while those for secondary modern schools decrease. This same difference can be seen within the school types. The spread of ability group reading score means in the 3-group schools increases between 11 and 16 in secondary modern schools but decreases in grammar and private schools. In contrast, the spread of mathematics scores decreases more in secondary modern schools than it does in grammar or private schools (see Appendix Tables A1 though A4).

As a result of these differing patterns for the two subjects and in the different kinds of schools and levels of ability groups, there is no general support for either of the derived hypotheses. Also, this means that any simple comparison of raw means across the age periods looks very different in the two subjects and for the different school types and ability groups.

The regression analyses also reflect these differences by subject, school type and ability group, although they also tend to obscure some of them. The significant effect of going to a grammar or private school on mathematics performance shown in Table 6 undoubtedly reflects, in part, the fact that the variance of mathematics scores increases in those schools while it decreases in secondary modern schools.

The regression analyses also suggest another contrast between the longitudinal effects in the two subject areas. The coefficients for ability groups in Table 6 are strikingly different for the two subject areas, even though both sets provide support for the basic hypothesis being investigated. The most striking difference is in the remedial group coefficients. They show a strong negative effect for reading but essentially no effect for mathematics. The other coefficients are also consistent with that effect. While being in a low ability group has a strong negative effect in reading, there is a less consistent effect for mathematics. Similarly, although the positive effect of being

in a high ability group is quite strong in mathematics, it is less so in reading. Although there is a clear contrast between high and low ability groups in the two subjects, the contrast is primarily due to high ability groups gaining in mathematics and low ability groups losing in reading. Since these losses and gains are in comparison with students in the same school types who were not in ability groups, this suggests that ability grouping serves different functions in the two subject areas. In mathematics it seems to serve the function of accelerating the high ability students; in reading it seems to serve the function of "setting aside" students in the low ability groups. While this may be an over-interpretation of relatively small differences, it is worth noting that the pattern just described is not only found for the full sample data reported in Table 6 but also for students in high and low ability groups within both secondary modern and comprehensive schools (see Table A6).

Finally, it may be important to return to the fact that the present analysis was conducted with British data and to pose the question of how these results compare with American analyses as well as the question of the relevance of these results to the American educational system.

As noted earlier, the most recent and most comprehensive empirical examination of the effects of ability grouping was conducted by Hotchkiss (1984). As has been the case with other American analyses, his analysis involved the differentiation of two ability groups - students in the college preparatory high school program and all other students. Hotchkiss refers to these students as being in or not being in the "academic track." One of the reasons for this simple distinction is that American high schools do not use a uniform system of differentiation, and the only generally recognized division is the one he used.

His analysis actually included a larger set of dependent variables than the present one, ten in all. They included not only test scores (verbal, math, science and civics) but also a number of measures of ambition, self esteem,

department, etc. The verbal and math test score analysis paralleled the present analysis in many ways. The dependent variables were measured in twelfth grade, and parallel measures were used from tenth grade. The analysis consisted of regressing the dependent variables on the earlier parallel measures together with an extended list of control variables.

The results of the analysis parallel those presented here in that the effects of being in the academic track were statistically significant but were reduced appreciably when a large number of control variables were included. He was able to go beyond the kind of analysis presented here, however, by including in his analysis a set of variables indicating the number of courses the students had had in various fields between when the time one and time two tests were administered. These measures do reduce the effects of academic track as expected, although there is still a significant effect of academic track on both verbal and math scores.

Hotchkiss seems to be impressed with the fact that, although the effect of academic track is statistically significant and cannot be explained away through reference to antecedent measures, the effect is quite small once these other factors are taken into account. One might be equally impressed with the fact that, even after including the set of courses taken, the tracking effect remains at all. Of course, this becomes a matter of whether one is impressed with a half full or a half empty glass.

A more important question to raise here might be how the results of the present analysis should be interpreted, in light of the kind of American analyses that have been conducted. A comparison with Hotchkiss' analysis may provide a basis for such an interpretation.

There are several basic differences between the Hotchkiss analysis and the present one:

(1) The position of students in ability groups used in the present analysis was determined by information provided by the student's teacher while that in the Hotchkiss analysis was provided by the student. Rosenbaum (1980) has reported great differences in the classification of students using those two methods, and he reports a stronger track effect when the information comes from the school.

(2) The Hotchkiss analysis, as well as other American analyses, have necessarily used only a two-track classification system while the present one has been able to use a much more complex system.

(3) It has been possible in the present analysis to include not only ability groups within the full sample but also a set of four school types and the ability groups within each type.

(4) The present analysis has included antecedent measures covering a broader period of the students' lives, ranging from age 7 to age 11 whereas the Hotchkiss analysis uses measures covering only a two-year period.

These several differences suggest that the present analysis is a more stringent test of the hypothesis of ability group effects than the previous analyses have been. While the above comparison uses Hotchkiss' analysis as the point of reference, almost all of the contrasts noted would have been the same whatever study was used for comparison.

But does the present analysis have relevance to the American case? This may be a more debatable question. It can be argued, certainly, that the British division between grammar and secondary modern schools has no counterpart in this country and thus it is pointless to emphasize that contrast for our purposes.

At the same time, it is important to remember that all of the findings of differences across ability groups are strongly upheld within the British comprehensive schools, which are as close as one can come in Great Britain to a counterpart of the American high school.

Thus, the more pressing question is whether the fact that the British schools report such an elaborate set of tracks, compared with the crude dichotomy in the American case, precludes any relevance of the present analysis for our purposes. I would argue that, on the contrary, the present analysis should alert us to the possibility that we have yet to conduct a wholly satisfactory analysis of the American system of tracking.

It is at least worth considering the possibility that, if we were to administer a questionnaire to American schools comparable to the one used in the NCDS, we might find that many (most?) American schools also track English and mathematics classes in much the same way reported by the British school officials. By depending on the global curriculum classification, we not only combine curricula which are undoubtedly different from each other (e.g., general and vocational), but we also obscure any variation within the academic track. If nothing else, it leads us to assume that all students in the academic track are at the same level in both English and mathematics. It seems unlikely that that is generally the case here, and the frequencies in various ability groups in the present analysis makes it clear it is not the case in Great Britain.

Thus, the present analysis may serve as a warning that the studies of American school organization have thus far been overly crude and have not attended to important differentiations that actually exist within our schools. The great concern for mathematics in our schools undoubtedly leads to a greater differentiation of ability groups in that subject than in English, and the British data analyzed here indicate that that is the case in Great Britain. As

a result, although the Hotchkiss analysis does seem to be as good as any previously conducted with American data, it can be argued that we have not yet adequately studied the effects of ability grouping with any American data.

REFERENCES

Alexander, K. L., M. Cook and E. McDill

1978 "Curriculum Tracking and Educational Stratification," American Sociological Review 43:47-66.

Alexander, K. L., and B. K. Eckland

1975 "Contextual Effects in the High School Attainment Process," American Sociological Review 40:402-16.

Alexander, K. L., and E. L. McDill

1976 "Selection and Allocation within Schools: Some Causes and Consequences of Curriculum Placement," American Sociological Review 41:963-80.

Boudon, R.

1974 Education, Opportunity and Social Inequality: Changing Prospects in Western Society. New York: Wiley.

Coleman, J. S., et al.

1966 Equality of Educational Opportunity. Washington, D.C.: U.S. Government Printing Office.

Davie, R., N. Butler and H. Goldstein

1972 From Birth to Seven. London: Longman.

Douglas, J. W. B., J. M. Ross and H. R. Simpson

1968 All Our Future. London: Peter Davies.

Esposito, D.

1973 "Homogeneous and Heterogeneous Ability Grouping: Principal Findings and Implications for Evaluating and Designing More Effective Educational Environments," Review of Educational Research, 43:163-79.

Findlay, W., and M. Bryan

1975 The Pros and Cons of Ability Grouping. Bloomington, Ind.: Phi Delta Kappa Educational Foundation.

Goldstein, H.

1976 "A Study of the Response Rates of Sixteen-Year-Olds in the National Child Development Study." Pp. 63-70 in K. Fogelman, Britain's Sixteen-Year-Olds. London: Macmillan.

Grant, W. V., and L. G. Eiden

1982 Digest of Education Statistics. Washington, D.C.: National Center for Education Statistics.

Heyns, B.

1974 "Social Selection and Stratification within Schools," American Journal of Sociology 79:1434:51.

Hotchkiss, L.

1984 Effects of Schooling on Cognitive, Attitudinal, and Behavioral Outcomes. Columbus, Ohio: The National Center for Research in Vocational Education.

Jencks, C. S., et al.

1972 Inequality: A Reassessment of the Effect of Family and Schooling in America. New York: Basic Books.

Oakes, J.

1982 "The Reproduction of Inequality: The Content of Secondary School Tracking," The Urban Review 14:107-20.

Rehberg, R. A., and E. R. Rosenthal

1978 Class and Merit in the American High School: An Assessment of the Revisionist and Meritocratic Arguments. New York: Longman.

Rosenbaum, J. E.

1976 Making Inequality: The Hidden Curriculum of High School Tracking. New York: John Wiley.

Rosenbaum, J. E.

1980 "Track Misperceptions and Frustrated College Plans: An Analysis of the Effects of Tracks and Track Perceptions in the National Longitudinal Survey," Sociology of Education 53:74-88.

Wohlwill, J. F.

1973 The Study of Behavioral Development. New York: Academic Press.

TABLE 1

FREQUENCIES, MEANS AND STANDARD DEVIATIONS OF TEST SCORES AT 11 AND 16, BY SCHOOL TYPES

		AGE 11		AGE 16	
	N	Mean	SD	Mean	SD
Boys Reading					
Total	4,797	16.41	6.36	25.88	6.88
Comprehensive	2,927	15.59	6.08	25.08	6.85
Grammar	518	22.42	4.58	31.37	3.23
Secondary Modern	1,070	14.20	5.41	24.02	6.89
Private	279	22.42	5.29	31.25	3.71
Girls Reading					
Total	4,602	16.46	5.94	25.65	6.50
Comprehensive	2,746	15.63	5.71	24.76	6.50
Grammar	609	21.77	4.15	31.09	2.81
Secondary Modern	986	14.18	5.06	23.54	6.32
Private	261	21.46	5.06	30.39	3.92
Boys Mathematics					
Total	4,768	17.72	10.46	13.77	7.19
Comprehensive	2,911	16.31	9.89	12.71	6.80
Grammar	513	29.37	6.08	21.46	5.36
Secondary Modern	1,071	13.48	8.41	11.07	5.41
Private	273	27.58	8.82	21.18	5.92
Girls Mathematics					
Total	4,572	17.18	10.01	12.35	6.65
Comprehensive	2,743	15.72	9.38	11.31	6.04
Grammar	600	27.68	6.54	19.75	5.05
Secondary Modern	976	12.77	8.08	9.31	4.88
Private	253	25.23	8.89	17.89	6.59

TABLE 2

FREQUENCIES, MEANS AND STANDARD DEVIATIONS OF TEST SCORES AT 11 AND 16, BY ABILITY GROUPS

		AGE 11		AGE 16	
	N	Mean	SD	Mean	SD
Boys Reading					
Not Grouped	1,007	17.19	6.63	26.48	6.72
Remedial	153	8.47	3.92	13.76	6.23
3-Groups - High	1,213	20.67	5.09	30.45	3.73
- Med	1,285	15.07	5.02	25.34	5.37
- Low	510	11.39	5.17	19.72	7.17
2-Groups - High	316	18.65	5.35	28.89	4.47
- Low	310	12.70	5.88	21.40	6.23
Girls Reading					
Not Grouped	1,149	17.47	6.04	26.62	6.25
Remedial	103	8.14	3.94	13.81	5.97
3-Groups - High	1,281	19.84	4.83	29.54	3.86
- Med	1,138	14.75	4.66	24.38	5.08
- Low	355	10.53	4.68	17.60	6.30
2-Groups - High	336	17.63	5.04	27.50	4.75
- Low	240	12.42	4.54	20.76	6.28
Boys Mathematics					
Not Grouped	421	17.86	10.59	13.42	7.26
Remedial	149	4.44	4.56	5.87	3.16
3-Groups - High	1,337	26.24	8.04	20.08	5.85
- Med	1,535	15.61	8.35	12.04	5.43
- Low	657	9.19	7.55	7.92	4.30
2-Groups - High	343	21.97	8.29	16.54	6.01
- Low	326	11.37	7.79	8.95	4.91
Girls Mathematics					
Not Grouped	417	15.35	9.49	10.59	6.15
Remedial	118	4.61	3.89	6.07	3.20
3-Groups - High	1,254	25.19	8.12	17.97	6.20
- Med	1,573	15.63	8.19	10.70	4.93
- Low	560	9.09	7.52	7.72	4.64
2-Groups - High	343	20.77	8.22	14.78	6.02
- Low	307	10.50	7.34	8.43	4.38

TABLE 3

MEAN CHANGES IN READING AND MATHEMATICS SCORES
BETWEEN 11 AND 16, WITHIN SCHOOL TYPES AND ABILITY GROUPS

	READING		MATHEMATICS	
	Boys	Girls	Boys	Girls
Mean Change, Total Sample	+9.47	+9.19	-3.95	-4.83
Deviations from Total for:				
Comprehensive School	+0.02	-0.06	+0.35	+0.42
Grammar School	-0.52	+0.13	-3.96	-3.10
Secondary Modern School	+0.35	+0.17	+1.54	+1.37
Private School	-0.64	-0.36	-2.45	-2.51
Not Grouped	-0.18	-0.04	-0.49	+0.07
Remedial Group	-4.18	-3.52	+5.38	+6.29
3-Group, - High	+0.31	+0.51	-2.21	-2.39
- Med	+0.80	+0.44	+0.38	-0.14
- Low	-1.14	-2.16	+2.68	+3.46
2-Group, - High	+0.77	+0.68	-1.48	-1.16
- Low	-0.37	-0.83	+2.02	+2.76

TABLE 4

STANDARD DEVIATION CHANGES FOR READING AND MATHEMATICS SCORES BETWEEN 11 AND 16, WITHIN SCHOOL TYPES AND ABILITY GROUPS

	READING		MATHEMATICS	
	Boys	Girls	Boys	Girls
SD Change, Total Sample	+0.52	+0.56	-3.27	-3.36
Deviations from Total for:				
Comprehensive School	+0.25	+0.23	+0.18	+0.02
Grammar School	-1.87	-1.90	+2.55	+1.87
Secondary Modern School	+0.96	+0.70	+0.27	+0.16
Private	-2.10	-1.70	+0.37	+1.06
Not Grouped	-0.43	-0.35	+0.04	+0.02
Remedial	+1.79	+1.47	+1.87	+2.67
3-Groups - High	-1.88	-1.53	+1.08	+1.44
- Med	-0.17	-0.14	+0.35	+0.10
- Low	+1.48	+1.06	+0.02	+0.48
2-Groups - High	-1.40	-0.85	+0.99	+1.16
- Low	+0.63	+1.18	+0.39	+0.40

TABLE 5

REGRESSION ANALYSIS OF AGE 16 TEST SCORES WITH SCHOOL TYPES AND ABILITY GROUPS

	Boys Reading	Girls Reading	Boys Math	Girls Math
School Type				
Grammar	-.243*	.397[a]	2.178[d]	3.099[d]
Secondary Modern	.279*	-.005*	-.309*	-.706[d]
Private	.110*	.450*	2.707[d]	2.367[d]
Adjusted R^2	.655	.679	.640	.607
R^2 Without Types	.654	.679	.628	.579
Ability Group				
Remedial	-5.659[d]	-4.523[d]	-1.072[b]	.538*
3-Group - High	.975[d]	.781[d]	2.741[d]	3.010[d]
- Med	.461[c]	-.173*	-.523[b]	-.156*
- Low	-2.022[d]	-3.241[d]	-1.507[d]	-.105*
2-Group - High	.890[c]	.482[a]	1.031[c]	1.653[d]
- Low	-1.405[d]	-1.744[d]	-1.563[d]	-.071
Adjusted R^2	.685	.705	.660	.610
R^2 Without Groups	.654	.679	.628	.579

* = non-significant
[a] = .05
[b] = .01
[c] = .001
[d] = .0001

TABLE 6

REGRESSION ANALYSIS OF AGE 16 TEST SCORES WITH SCHOOL TYPES AND ABILITY GROUPS, WITH CONTROLS

	Boys Reading	Girls Reading	Boys Math	Girls Math
School Type				
Grammar	-.622[b]	-.007*	1.812[d]	2.521[d]
Secondary Modern	.362[a]	.088*	-.267*	-.658[d]
Private	-.265*	.170*	1.770[d]	1.470[d]
Adjusted R^2	.672	.700	.655	.626
R^2 Without Types	.654	.679	.628	.579
Ability Group				
Remedial	-5.006[d]	-3.643[d]	-.699*	1.170[a]
3-Group - High	.908[d]	.827[d]	2.376[d]	2.678[d]
- Med	.538[b]	.020*	-.627[b]	-.240*
- Low	-1.755[d]	-2.669[d]	-1.303[d]	.078*
2-Group - High	.714[b]	.510[a]	.693[a]	1.512[d]
- Low	-1.279[d]	-1.459[d]	-1.328[d]	.229*
Adjusted R^2	.694	.719	.674	.636
R^2 Without Groups	.685	.705	.660	.610

* = non-significant
[a] = .05
[b] = .01
[c] = .001
[d] = .0001

TABLE A1

BOY'S FREQUENCIES, MEANS AND STANDARD DEVIATIONS OF AGE 11 AND 16 READING SCORES, BY ABILITY GROUP WITHIN SCHOOL TYPES

		AGE 11		AGE 16	
	N	Mean	SD	Mean	SD
TOTAL	4,797	16.41	6.36	25.88	6.88
Comprehensive	2,927	15.59	6.08	25.08	6.85
Not Grouped	435	15.03	6.20	24.42	6.87
Remedial	107	8.70	3.76	13.69	5.74
3-Group - High	748	20.34	4.98	30.25	3.78
- Med	868	14.69	4.87	25.01	5.27
- Low	340	11.16	4.83	19.09	6.56
2-Group - High	199	18.00	5.06	28.54	4.38
- Low	230	12.21	4.56	20.97	5.83
Grammar	518	22.42	4.58	31.37	3.23
Not Grouped	224	22.25	4.35	31.03	3.68
Remedial	*	*	*	*	*
3-Group - High	177	23.12	4.72	32.03	2.51
- Med	39	21.31	4.32	31.21	2.82
- Low	27	19.81	4.21	30.00	2.70
2-Group - High	38	23.63	5.00	32.18	2.60
- Low	13	21.23	4.28	29.23	4.78
Secondary Modern	1,070	14.20	5.41	24.02	6.89
Not Grouped	244	13.95	5.06	24.07	6.29
Remedial	45	7.82	4.24	13.62	7.13
3-Group - High	163	17.53	4.29	28.66	4.32
- Med	360	15.01	4.81	25.21	5.41
- Low	132	9.86	4.41	18.35	7.15
2-Group - High	61	17.41	4.74	27.56	4.64
- Low	65	12.55	5.48	21.06	6.68
Private	279	22.42	5.29	31.25	3.71
Not Grouped	104	22.89	5.76	30.99	4.63
Remedial	*	*	*	*	*
3-Group - High	125	23.36	4.44	31.78	2.61
- Med	18	21.39	4.67	30.78	2.46
- Low	11	16.27	4.22	30.36	4.06
2-Group - High	18	19.56	5.66	30.22	4.80
- Low	*	*	*	*	*

*Means and standard deviations are not computed if the category frequency is less than 10.

TABLE A2

GIRL'S FREQUENCIES, MEANS AND STANDARD DEVIATIONS OF AGE 11 AND 16 READING SCORES, BY ABILITY GROUP WITHIN SCHOOL TYPES

		AGE 11		AGE 16	
	N	Mean	SD	Mean	SD
TOTAL	4,602	16.46	5.94	25.65	6.50
Comprehensive	2,746	15.63	5.71	24.76	6.50
Not Grouped	433	14.77	5.32	23.80	6.35
Remedial	70	7.89	3.86	13.03	5.78
3-Group - High	893	19.47	4.85	29.21	4.00
- Med	744	14.31	4.41	23.98	4.91
- Low	239	10.45	4.39	17.30	5.78
2-Group - High	209	17.08	5.20	26.74	5.08
- Low	158	11.85	3.85	19.77	5.38
Grammar	609	21.77	4.15	31.09	2.81
Not Grouped	321	22.07	4.09	31.29	2.62
Remedial	*	*	*	*	*
3-Group - High	171	22.42	3.82	31.67	2.20
- Med	52	19.79	4.85	29.37	4.01
- Low	12	17.75	4.49	27.83	4.57
2-Group - High	28	21.75	3.89	31.07	2.43
- Low	25	19.64	2.58	29.80	2.68
Secondary Modern	986	14.18	5.06	23.54	6.32
Not Grouped	236	13.58	4.34	23.00	5.45
Remedial	31	8.06	3.18	14.81	5.54
3-Group - High	160	18.01	4.31	28.31	3.84
- Med	322	14.60	4.50	24.20	5.16
- Low	96	9.46	4.16	16.36	6.03
2-Group - High	86	16.97	4.12	27.64	3.93
- Low	55	10.47	3.74	19.09	6.22
Private	261	21.46	5.06	30.39	3.92
Not Grouped	159	21.36	5.03	30.21	4.12
Remedial	*	*	*	*	*
3-Group - High	57	22.93	4.49	31.74	2.31
- Med	20	20.65	4.42	29.50	2.93
- Low	*	*	*	*	*
2-Group - High	13	22.08	4.23	31.08	2.36
- Low	*	*	*	*	*

*Means and standard deviations are not computed if the category frequency is less than 10.

TABLE A3

BOY'S FREQUENCIES, MEANS AND STANDARD DEVIATIONS OF AGE 11 AND 16 MATHEMATICS SCORES, BY ABILITY GROUP WITHIN SCHOOL TYPES

		AGE 11		AGE 16	
	N	Mean	SD	Mean	SD
TOTAL	4,768	17.72	10.46	13.77	7.19
Comprehensive	2,911	16.31	9.89	12.71	6.80
Not Grouped	165	14.88	9.23	10.88	5.65
Remedial	111	4.22	3.55	5.59	2.99
3-Group - High	764	25.37	7.93	19.53	5.80
- Med	997	14.67	7.74	11.35	4.93
- Low	428	8.63	6.58	7.38	3.61
2-Group - High	200	21.19	7.80	15.66	5.93
- Low	246	10.59	6.71	8.39	4.07
Grammar	513	29.37	6.08	21.46	5.36
Not Grouped	83	29.83	5.33	21.80	5.42
Remedial	*	*	*	*	*
3-Group - High	229	30.86	5.64	22.93	4.82
- Med	103	27.08	6.70	19.57	5.82
- Low	26	24.42	6.70	18.35	4.44
2-Group - High	52	29.71	5.05	20.33	5.40
- Low	19	27.58	5.42	19.95	4.27
Secondary Modern	1,071	13.48	8.41	11.07	5.41
Not Grouped	144	12.68	8.24	10.04	4.97
Remedial	36	3.89	4.20	6.08	2.35
3-Group - High	186	20.42	7.06	16.18	5.24
- Med	396	14.23	7.44	11.07	4.32
- Low	183	6.93	5.40	6.96	2.89
2-Group - High	70	18.13	7.45	15.46	5.18
- Low	56	8.96	6.02	7.80	4.00
Private	273	27.58	8.82	21.18	5.92
Not Grouped	29	26.31	8.02	20.59	5.79
Remedial	*	*	*	*	*
3-Group - High	158	30.58	6.77	23.20	4.52
- Med	39	23.54	9.25	19.79	5.66
- Low	20	21.95	10.63	14.75	5.31
2-Group - High	21	22.95	10.03	19.14	6.62
- Low	*	*	*	*	*

*Means and standard deviations are not computed if the category frequency is less than 10.

TABLE A4

GIRL'S FREQUENCIES, MEANS AND STANDARD DEVIATIONS OF AGE 11 AND 16 MATHEMATICS SCORES, BY ABILITY GROUP WITHIN SCHOOL TYPES

		AGE 11		AGE 16	
	N	Mean	SD	Mean	SD
TOTAL	4,572	17.18	10.01	12.35	6.65
Comprehensive	2,743	15.72	9.38	11.31	6.04
Not Grouped	176	13.04	8.63	8.86	5.26
Remedial	81	4.69	3.73	6.27	3.25
3-Group - High	709	23.71	7.99	16.81	6.01
- Med	1,003	14.53	7.62	9.99	4.22
- Low	356	8.00	6.00	7.09	3.54
2-Group - High	212	20.22	7.55	14.10	5.71
- Low	206	9.33	5.86	7.35	3.27
Grammar	600	27.68	6.54	19.75	5.05
Not Grouped	57	27.09	6.14	18.98	5.10
Remedial	*	*	*	*	*
3-Group - High	299	29.55	6.11	21.60	4.42
- Med	141	25.35	6.26	16.99	5.09
- Low	37	24.95	6.37	17.46	4.73
2-Group - High	39	28.10	6.14	20.49	4.52
- Low	27	23.56	7.49	17.30	3.44
Secondary Modern	976	12.77	8.08	9.31	4.88
Not Grouped	138	11.59	7.19	8.25	4.08
Remedial	34	3.97	3.36	5.41	3.03
3-Group - High	142	19.58	7.43	13.03	5.24
- Med	374	13.91	7.37	9.57	4.50
- Low	152	6.80	5.01	6.16	3.28
2-Group - High	69	17.14	8.09	12.64	5.10
- Low	67	8.00	5.19	7.81	3.38
Private	253	25.23	8.89	17.89	6.59
Not Grouped	46	20.93	8.27	13.80	5.52
Remedial	*	*	*	*	*
3-Group - High	104	30.48	6.20	22.24	5.07
- Med	55	22.65	7.61	15.40	5.08
- Low	15	19.27	10.24	14.47	5.67
2-Group - High	23	24.30	9.21	17.74	7.09
- Low	*	*	*	*	*

*Means and standard deviations are not computed if the category frequency is less than 10.

TABLE A5

EFFECTS OF ABILITY GROUP ASSIGNMENT ON AGE 16 TEST PERFORMANCE, CONTROLLING EARLIER TEST PERFORMANCE

	GROUPS						
	Remedial	3 High	3 Med	3 Low	2 High	2 Low	Adjusted R^2
Comprehensive							
Reading - Boys	-5.875[d]	1.256[d]	.621[c]	-2.148[d]	1.272[c]	-1.273[d]	.676
- Girls	-5.028[d]	1.174[d]	.193*	-3.215[d]	.624[a]	-1.747[d]	.694
Math - Boys	-1.314[b]	3.963[d]	-.130*	-1.531[d]	1.576[d]	-1.352[c]	.632
- Girls	.673*	3.774d	.299*	.081*	2.367[d]	-.172*	.565
Grammar							
Reading - Boys	-	.759[b]	.303*	-.448*	.644*	-1.423*	.221
- Girls	-	.257*	-1.114[c]	-1.735[b]	-.175*	-.615*	.400
Math - Boys	-	.536*	-.576*	.007*	-1.356*	-.633*	.294
- Girls	-	1.547[b]	-1.285[a]	-1.018*	.942*	-.233*	.427
Secondary Modern							
Reading - Boys	-4.291[d]	.805*	.170*	-1.992[d]	.055*	-1.503[b]	.641
- Girls	-3.119[d]	1.340[b]	.187*	-3.177[d]	1.610[b]	-1.775[b]	.600
Math - Boys	-.900*	3.276[d]	.170*	-1.391[c]	3.228[d]	-1.438[a]	.522
- Girls	-.658*	2.243[d]	.492*	-.915[a]	2.565[d]	.199*	.352
Private							
Reading - Boys	-	.445*	.639*	2.734[b]	-.054*	4.089*	.493
- Girls	-1.414*	.789*	-.681*	.098*	.399*	2.478*	.633
Math - Boys	-2.695*	.893*	.690*	-3.195[c]	-.477*	-8.310[d]	.629
- Girls	-1.225*	4.001[d]	.500*	-.029*	.969*	-2.063*	.583

* = Non-significant
[a] = .05
[B] = .01
[c] = .001
[d] = .0001

TABLE A6

EFFECTS OF ABILITY GROUP ASSIGNMENT ON AGE 16 TEST PERFORMANCE, CONTROLLING EARLIER TEST PERFORMANCE AND OTHER VARIABLES

		GROUPS						
		Remedial	3 High	3 Med	3 Low	2 High	2 Low	Adjusted R^2
Comprehensive								
Reading	- Boys	-5.097[d]	1.087[b]	.629[b]	-1.851[d]	.934[b]	-1.280[d]	.688
	- Girls	-4.048[d]	1.005[d]	.268*	-2.447[d]	.479*	-1.416[d]	.708
Math	- Boys	-.813*	3.622[d]	-.208*	-1.225[c]	1.295[b]	-1.035[b]	.645
	- Girls	1.151[a]	3.258[d]	.051*	.213*	2.031[d]	.026*	.585
Grammar								
Reading	- Boys	-	.827[a]	.661*	-.408*	.376*	-1.020*	.220
	- Girls	-	.387*	-.772[a]	-1.438[a]	-.041*	-.686*	.416
Math	- Boys	.746*	.350*	-.757*	-.813*	-1.827[a]	-.724*	.324
	- Girls	-	1.233[a]	-1.456[a]	-.948*	.635*	.010*	.456
Secondary Modern								
Reading	- Boys	-3.818[d]	.639*	.062*	-1.849[c]	.091*	-1.212[a]	.658
	- Girls	-2.183[a]	1.293[b]	.240*	-3.149[d]	1.439[c]	-1.729[a]	.618
Math	- Boys	-.886*	3.005[d]	.143*	-1.213[a]	2.932[d]	-1.370[a]	.513
	- Girls	-.373*	2.177[d]	.369*	-.852*	2.514[d]	.348*	.369
Private								
Reading	- Boys	-	.464*	.548*	2.535[a]	-.024*	6.914[a]	.524
	- Girls	1.217*	.927*	-.482*	-.564*	-.008*	3.102*	.608
Math	- Boys	-1.611*	.506	.143*	-4.144[b]	-2.113*	-8.646[d]	.626
	- Girls	-1.753*	3.393[b]	.331*	-.941*	.093*	-2.343*	.591

* = Non-significant
a = .05
b = .01
c = .001
d = .0001