Student Achievement and the Changing American Family
An Executive Summary

David W. Grissmer, Sheila Nataraj Kirby, Mark Berends, Stephanie Williamson
This research described in this report was supported by a grant from the Lilly Endowment Inc.

ISBN: 0-8330-1609-1

RAND is a nonprofit institution that helps improve public policy through research and analysis. RAND's publications do not necessarily reflect the opinions or policies of its research sponsors.

Published 1994 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
To order RAND documents or to obtain additional information, contact Distribution Services: Telephone: (310) 451-7002; Fax: (310) 451-6915; Internet: order@rand.org.
Errata Sheet


p. 13 At the end of the first paragraph, insert the sentence: “These findings are not unique to this study but have been reported previously by several other researchers.”

p. 20 Second complete paragraph, first sentence, should read: “As pointed out by others, using the SAT scores . . . .”

p. 22 First sentence after the heading Explaining minority test score gains should read: “Minorities have made significant gains in test scores over the last 20 years and our study shows that a large proportion of these gains was unexplained by family changes.”
Student Achievement and the Changing American Family
An Executive Summary

David W. Grissmer, Sheila Nataraj Kirby, Mark Berends, Stephanie Williamson

Supported by the Lilly Endowment Inc.
There is considerable debate in our society and in the research community about the direction and causes of changes in student performance over the last 25 years: whether student performance is getting better or worse, whether the dramatic changes in family and racial/ethnic characteristics have affected average student achievement, and whether the greatly expanded investment in education and other social programs and policies directed toward equal educational opportunities were effective in improving student performance.

These questions are among the most important public policy issues affecting our society's future. The proficiencies and future performance of our children will be partly responsible for our competitive economic position in the world economy. In addition, children's outcomes will partly determine how much future public spending will be required to pay for such programs as prisons and the criminal justice system, welfare, unemployment, and job training, as well as health expenditures arising from treatment of addictions and victims of violent crimes.

Better answers to these questions would help determine how to more effectively allocate the approximately $275–325 billion of public resources for K–12 education and social programs directed to support families and improve student outcomes. Answers to these basic questions would also help establish whether fundamental school reform is needed and, if so, help provide directions for reform of schools and school financing.
The effectiveness of public policies and investments directed toward children cannot be evaluated without accounting for the changing characteristics and demographics of the American family. Many believe that the family has deteriorated in its capacity to support the development of children. They point to rising levels of single-parent families and working mothers, a greater proportion of children in poverty households, and a higher incidence of births among young, unmarried mothers.

Besides changes in family characteristics, the proportion of children from Hispanic and Asian backgrounds has increased markedly over the last 25 years, mainly due to immigration. There has also been a more gradual increase in the proportion of children who are black. These changing demographic trends also can affect average student performance. If family and demographic characteristics are key factors related to changes in average student achievement, then they must be taken into account before attempting to evaluate the effectiveness of increases in public investments and changing public policies.

In this study, we first estimate the effect that changing family characteristics and race/ethnicity of students would be expected to have on national mathematics and verbal/reading achievement score trends of 14–18-year-old youth between 1970/1975 and 1990. Second, we compare these estimated effects from changing family/demographic characteristics to actual trends in national achievement scores to see how much of the actual trend might be accounted for by changing families and demographics. Third, we estimate the residual effect that cannot be accounted for by family/demographic trends and offer several hypotheses that might help explain the patterns of the residual effects.

This report is intended for government policymakers, educators, researchers, parents, and taxpayers interested in how to effectively and efficiently foster higher student achievement. Research support to build the database used in this analysis was provided jointly by RAND’s Institute on Education and Training (IET) through a grant from the Lilly Endowment Inc. and by the Office of the Assistant Secretary of Defense (P&R). The Department of Defense is the nation’s largest employer of youth, and as such is vitally concerned with changing levels of achievement among the nation’s youth. The re-
sults and policy implications of our research are being documented in two separate reports: one for those primarily concerned with domestic issues, and the other focusing mainly on concerns of defense policymakers. The results for domestic audiences will be presented in a forthcoming report, MR-488-LE, Student Achievement and the Changing American Family; the present document summarizes our results. This report was written under the aegis of the IET.
CONTENTS

Preface .......................................................... iii
Figures .......................................................... ix
Acknowledgments ............................................... xi

BACKGROUND AND PURPOSE OF THE STUDY ............. 1

METHODOLOGY .................................................. 4
   Step 1: Estimating How Much Family and Demographic
          Characteristics Affect Test Scores ....................... 4
   Step 2: Estimating How Much Changing Families and
          Change Test Scores ......................................... 5
   Step 3: Estimating How Much Changing Family and
          Demographic Characteristics Can Account for Actual
          Test Score Changes ........................................ 5

RESULTS ........................................................ 6
   How Much Do Family and Demographic Characteristics
   Affect Test Scores? ............................................. 7
   How Much Would Changing Families and Demographics
   Change Test Scores? .......................................... 10
   How Much of Test Score Changes Can Be Accounted for
   by Changes in Family and Demographics? ............... 13

DISCUSSION OF RESULTS AND FUTURE RESEARCH ...... 15
   The American Family ........................................ 15
   Test Scores as Indicators of School Quality ............... 19
FIGURES

1. Simple Differences in Mean Mathematics Test Scores for Selected Groups, NLSY and NELS ............... 7
2. Net Differences in Mean Mathematics Test Scores for Selected Groups, NLSY and NELS ............... 9
3. Estimated Family and Demographic Effects on Mathematics Test Scores Between 1975 and 1990 by Racial/Ethnic Group ........................................ 12
4. Estimated Family and Demographic Effects on Verbal Test Scores Between 1975 and 1990 by Racial/Ethnic Group ........................................ 13
8. Comparison of Differences Between Racial/Ethnic Groups in NAEP Verbal Scores in 1975 and 1990 .... 17
ACKNOWLEDGMENTS

We are grateful to the Lilly Endowment Inc. for their support of the Institute for Education and Training at RAND, under whose aegis this work was conducted. We are also grateful to Stephen Sellman, Director of Accession Policy, Office of the Secretary of Defense, for his support and recognition of the commonality of defense and domestic issues with respect to future student achievement. Their support and the support of Georges Vernez and Roger Benjamin, directors of the IET, allowed us to examine these broad and fundamental questions about public policy, the American family, and the educational system.

We are indebted to Paul Hill of RAND and Richard Murnane of Harvard University for their insightful reviews and useful suggestions for improving this report. The report has benefited by comments from Bernard Rostker, director of RAND's Defense Manpower Research Center. We also appreciate the work of Robert Young, who assisted in building the data files, and Luetta Pope and Susan Spindel, who contributed to typing the final document.
BACKGROUND AND PURPOSE OF THE STUDY

There is a continuing national debate on the quality of our children’s family environment, the quality of their schools, and how changes in families and schools may be affecting the level of student achievement. Within this debate, questions remain about whether public policies and increased investments in education and social programs are effective in improving student achievement. In addition, as the student population becomes more racially and ethnically diverse, there is growing concern about the inequality of educational outcomes between minority and nonminority students.

Motivating these concerns are negative perceptions about achievement trends, the changing family environment, and the effectiveness of social programs and public education. Specifically, these include the following:

- A perceived decline in student achievement as measured by scores on the widely publicized Scholastic Aptitude Test (SAT).

- A perceived deterioration in the family environment, with particular emphasis on four trends: increase in the number of teen mothers and out-of-wedlock births, increase in the number of children living in poverty, increase in the proportion of mothers working, and increase in the number of children living in single-parent families.

- The perceived ineffectiveness of the very large increases in the real per-pupil K-12 educational expenditures and other social
programs over the last 25 years in producing higher student achievement.

Trying to sort out the relative contributions of families, schools, and social programs to student achievement is a complex exercise, for several reasons. First, there are conflicting trends in student achievement, depending on which tests are used, and a great deal of caution needs to be observed in selecting a representative test and in interpreting its results. Second, explaining trends is difficult because several factors perceived to affect student achievement have all changed dramatically: the family environment, demographic mix of students, school quality, public policies directed toward providing equal educational opportunity, and public investment in schools and social programs. Third, assessing the effect of public policies and investment is problematic partly because empirical evidence indicates that family and demographic changes probably have the largest effects on test scores; thus, family/demographic effects on student achievement need to be estimated before making assessments of the effect of public policies and investment.

These substantial changes—in family and demographic characteristics of students, the initiatives to provide equal educational opportunity through such policies as integration of public schools and bilingual education programs, and the increased public investment in schools and social programs—all combine to make the last 25 years a unique period in our history and provide a unique opportunity to understand the trends in student achievement and their causes. Understanding how our families and schools have changed, the impact of these changes on student performance, and whether public policies and investment make a difference will help provide answers to some of the most important public policy questions affecting the future of our society.

Answering such questions and sorting out the relative contributions of the various factors to student achievement is the main purpose of this overall project. The present study primarily focuses on estimating the change in achievement test scores that can be attributed to
changing family and demographic characteristics.\textsuperscript{1} Family characteristics included in the analysis are family income, family size, parental education levels, age of mother at birth, working mothers, and single-parent families. Our analysis estimates the expected effects that changing family environment and demographic characteristics would have had on the student achievement scores of a national sample of students ages 14–18 between 1970/1975 and 1990. The direction of the predicted effects of these family changes on test scores can provide evidence on whether the family environment in 1990 is more or less supportive of student achievement compared with the environment of similarly aged students in 1970/1975. If we predict a decline in scores due to family changes, this would reinforce the common perception that the family environment has indeed deteriorated over this time period; a predicted increase in scores would indicate more supportive family characteristics and environment in 1990.

We compare our predicted changes in test scores from family/demographic changes to actual changes in achievement scores of national samples of youth from 1970/1975 to 1990 in order to gauge the effect of factors other than family and demographic characteristics on student achievement. If the residual between actual changes in test scores and predicted changes based on family/demographic effects alone is positive, then this would suggest that other factors had a positive effect on test scores, while a negative residual would suggest the opposite.

Two major factors that could help account for a positive residual between actual score changes and those accounted for by family/demographic changes are (a) changing public policies in the area of equal educational opportunity and increased levels of public investment in schools and children, and (b) changing productivity of schools. The residual can provide some evidence about whether effects from public policy and public investment and changing schools are present.

\textsuperscript{1}The summary presented here is based on analyses documented in a detailed technical report. See David W. Grissmer, Sheila Nataraj Kirby, Mark Berends, and Stephanie Williamson, \textit{Student Achievement and the Changing American Family}, Santa Monica, CA: RAND, MR-488-LE, 1994.
To help determine if the pattern of residual differences is consistent or inconsistent with positive effects from public policies, public investment, and schools, we estimate family effects and residual differences for black, Hispanic, and non-Hispanic white youth. It is possible that minority and nonminority families may have changed in different ways over this period and that the predicted family effects differ across these groups. In addition, the effects of public policies and investment would not be the same across racial/ethnic groups because public policies concerning equal opportunity and additional public investment in education and social programs were differentially targeted toward minority and/or lower-income families, children, and school districts. Thus, we might expect the indicators of these effects to be larger for minority groups. As such, our analysis provides separate estimates of family effects, the predicted change in test scores based on these family effects, and the gap between actual and predicted test score changes for blacks, Hispanics, and non-Hispanic whites.

METHODOLOGY

The methodology consists of three steps: (1) developing quantitative models linking student achievement to family and demographic characteristics; (2) using these models to predict test scores for a national sample of children using their family and demographic characteristics from 1970, 1975, and 1990; and (3) comparing the mean changes in these predicted test scores between 1970/1975 to 1990 (changes due to family and demographic characteristics) to actual changes in test scores of a national sample of children and estimating a residual not accounted for by family/demographic factors. This pattern of residuals can provide initial evidence that additional factors—hypothesized to be primarily changes in schools, public policies, and public investment—may have affected student achievement.

Step 1: Estimating How Much Family and Demographic Characteristics Affect Test Scores

We first estimate models linking test scores to family and demographic characteristics; we have used two quite different nationally representative samples of adolescents. The data sets are the National
Longitudinal Survey of Youth (NLSY), 1980, from which we selected students aged 15–18 years, and the National Education Longitudinal Survey (NELS), 1988, which samples eighth graders. The dependent variables in the models are the standardized scores for mathematics and verbal/reading tests that were administered to all children in the samples. Test scores are assumed to be a function of a set of independent family and demographic variables which are common to both surveys. These include family income, family structure (single-parent or two-parent households), family size, parental education, labor-force participation of the mother, age of the mother at child’s birth, and race/ethnicity. These are the primary family variables that have changed over the last 20 years and have been linked to student achievement.


These equations are then used to predict test scores at the individual level for a representative sample of U.S. children of similar ages in 1970, 1975, and 1990, extracted from the March Current Population Surveys. We compute the mean shift of the distribution of test scores from 1970/1975 to 1990 to provide an overall measure of the net effect of changing family and demographic characteristics on test scores. We make black/nonblack comparisons using the 1970 and 1990 data; comparisons for three racial/ethnic groups—non-Hispanic whites, Hispanics, and blacks—are done using the 1975 and 1990 predicted test scores.²

**Step 3: Estimating How Much Changing Family and Demographic Characteristics Can Account for Actual Test Score Changes**

The third step is to compare the changes in test scores predicted from changes in family and demographic characteristics to actual changes in national test scores. Although Scholastic Aptitude Test

(SAT) scores are probably what the public most often uses to form opinions about national test score trends, these scores are seriously flawed as indicators of how the average achievement of American students is changing, for several reasons. First, the SAT sample is not a representative sample of U.S. students. Second, the sample contains a constantly changing proportion and composition of students, and this has introduced a downward bias in scores over time. Third, from our perspective, a more serious flaw is that the SAT sample excludes students not going to college. As evidence cited below indicates, the largest changes in scores over the last 20 years have probably occurred among lower-scoring students, who are less likely to go to college or to take the SAT. Thus, the SAT scores probably missed the students making the largest changes.

The National Assessment of Educational Progress (NAEP) is a set of standardized tests that has been given by the Department of Education since the early 1970s to a nationally representative sample of students aged 9, 13, and 17 years. The questions used for comparisons over time have not changed and thus can be used for making comparisons of student achievement over time. These tests provide the best data to monitor the achievement trends of U.S. students over the last 25 years. We compare the NLSY results for 15–18-year-old youth to NAEP scores for 17-year-olds and the NELS results for eighth graders to NAEP scores for 13-year-olds. These comparisons allow us to calculate what proportion of the actual score changes can be attributed to family/demographic changes and what remains to be explained by other factors.

RESULTS

The analysis and evidence reported here support a more positive picture than is usually drawn of the achievement of American students aged 14–18 years, the capacity of American families to support that achievement, and the effectiveness of public policies and public investment.

---

How Much Do Family and Demographic Characteristics Affect Test Scores?

The results from the NELS and NLSY both show large differences in test scores for family/demographic characteristics and great similarity in the direction and relative significance of these differences. Figure 1 shows simple comparisons of mathematics test scores among youth in different types of families from the NLSY and NELS.\(^4\)

---

\(^4\) The mathematics and verbal/reading test score differences reported in Grissmer et al. (1994) show fairly similar patterns and sizes of differences.

\(^5\) We utilize a consistent measure—proportion of a standard deviation—throughout to measure differences in test scores. Another measure commonly used in reporting test scores is the percentile. This shows the relative standing of a particular score and measures the proportion of children scoring lower than that score. A 0.10 of a standard deviation difference in test scores is approximately 3.4 percentile points for most children. So two groups of children whose average scores differ by 0.10 of a standard deviation would indicate that one group scores—on average—3.4 percentile points higher than the other group.
The figure shows large differences among the average test scores of children living in families with different levels of parental education or of different racial/ethnic background. For instance, a child whose mother or father graduated from college scores approximately 1.0 standard deviation higher than a child whose mother or father did not graduate from high school, while black and Hispanic youth score from 0.50 to 1.0 of a standard deviation lower than non-Hispanic white youth.

Somewhat smaller test score differences are evident among young people living in families with different levels of annual income ($40,000 versus $15,000), families of different size (four siblings versus one sibling), having older versus younger mothers (age 30 at birth versus age 18) and living in two-parent versus single-parent families. For instance, children living in two-parent families score about 0.30 to 0.40 of a standard deviation higher than those in single-parent families, while children in large families score approximately 0.30 of a standard deviation lower than children from smaller families. There is little difference in test scores between those with working versus nonworking mothers.

Public debate and the press often focus on these simple comparisons of achievement scores for different family and demographic characteristics and mistakenly attribute the difference in scores between two groups to the particular characteristic in which the groups differ. Such inferences are misleading, however, because the students being compared usually differ in several characteristics, not just the one being cited. For instance, young people in higher-income families are also more likely to have parents with higher levels of education and to be nonminority. Thus, the difference in average test scores between children from high-income versus low-income families is probably due to a combination of factors, not just income alone. A better measure of the effect of income on test scores is a controlled comparison of two groups of young people who have similar family characteristics except for income. This is true for other characteristics as well.
Figure 2 summarizes these controlled comparison differences for mathematics scores.\(^6\)

This figure shows that the net effect of each factor is considerably smaller than the simple comparisons in Figure 1. However, the controlled differences remain significant for certain characteristics. For example, youth whose parents are college graduates score about 0.50 of a standard deviation higher than youth who are otherwise similar but have parents who did not graduate from high school. In addition, controlling for other family characteristics, the difference between blacks and non-Hispanic whites is 0.50 of a standard deviation, and the difference between Hispanics and non-Hispanic whites is somewhat smaller. Youth with different levels of family income or

![Figure 2—Net Differences in Mean Mathematics Test Scores for Selected Groups, NLSY and NELS](image)

\(^6\)These effects are derived by using the estimates from our multivariate model of student achievement. Multivariate models allow us to examine the effect of a particular characteristic, holding constant other important variables.
different family sizes show much smaller differences in test scores. Controlled test score differences due to family structure and labor-force participation of the mother appear to be negligible. These results suggest that the simple differences between youth scores in single- and two-parent families arise directly from other differences in family characteristics, such as family income, parental education, or family size rather than the structure of the one-parent versus two-parent family itself.

How Much Would Changing Families and Demographics Change Test Scores?

We use the estimates from the multivariate models (which formed the basis for Figure 2 above) to predict the changes in test scores that would be expected due to the changes in family and demographic characteristics that occurred between 1970/1975 and 1990.

We find that 14–18-year-olds living in U.S. families in 1990 would be predicted to score higher, not lower, on tests compared to youth in families in 1970. The size of the shift in mean scores is approximately 0.20 of a standard deviation. This means that youth in 1990 would be expected to have scores about 7 percentile points higher than their counterparts in 1970, based on combined changes in demographic and family characteristics. It should be emphasized that these findings estimate average effects when taking account of all American families with 14–18-year-olds.

Our analysis suggests that the most important family influences on student test scores are the level of parental education, family size, family income, and the age of the mother when the child was born.

Of these variables, the two that have changed most dramatically in a favorable direction are parental education levels and family size. Children in 1990 are living with better-educated parents and in smaller families. These factors are the primary reasons that changes in family characteristics would predict higher test scores. For example, 7 percent of mothers of 15–18-year-old children in 1970 were college graduates, compared to 16 percent in 1990, whereas 38 percent did not have high school degrees in 1970, compared to only 17 percent in 1990. Similar, but somewhat smaller, changes occurred in the educational attainment of fathers. Changes in family size were
also dramatic. Only about 48 percent of 15–18-year-old children lived in families with at most one sibling in 1970, compared to 73 percent in 1990.

Our analysis indicates that average family income changed little over the period 1970 to 1990 (in real terms), so it would not be expected to affect average test scores. However, the decline in family size coupled with unchanged average family income means that family income per child actually increased from 1970 to 1990.

One change that has had a slight negative effect on test scores is the small decline in the average age of the mother at the child's birth. This is due partly to increased births to younger mothers, but also to the decline in family size, which reduces the number of children born to older mothers.

The effect of the large increase in working mothers and single-parent families is more complex (discussed in more detail below). Our equations imply that the large increase in working mothers would—other things equal—have a negligible or small positive effect on youth test scores. However, the mother's labor-force participation is measured when the youth was approximately 14 years old, so our results may not apply to younger children.

In the case of the increase in single mothers, our models imply no negative effects from the changed family structure alone. However, such families tend to have much lower income levels, so the predictions for youth in these families incorporate a negative impact due to increasing numbers of poor, single-parent families.7

We turn now to the results by racial/ethnic group between 1975 and 1990. Figures 3 (mathematics) and 4 (verbal) show the estimated family effects separately for non-Hispanic whites, blacks, and Hispanics as well as the total youth population between 1975 and 1990. Higher mathematics scores in 1990 would be expected for 14–15 and 15–18-year-olds for each racial/ethnic group based on changing family characteristics. The data show that non-Hispanic white and

---

7A more technical discussion of these complex effects is given in the main report, Grissmer et al. (1994), Chapter Five.
black youth have similar predicted family gains of approximately 0.20 of a standard deviation, but Hispanic youth show smaller gains of approximately 0.05 of a standard deviation. Verbal/reading score comparisons show slightly higher gains than for mathematics, although the pattern is similar by racial/ethnic groups. The positive changes in the black family in terms of increased parental education and reduced family size are actually greater than those for non-Hispanic white families, but there were offsetting increases in births to younger and single mothers. The smaller gains for Hispanic youth are explained by smaller increases in parental education, falling family income, and smaller reductions in family size compared to black families. This is probably due to the continuing immigration of large number of Hispanic families into the population, many of whom may have lower levels of educational achievement and fewer labor-market skills than previous waves of immigrants.8

Figure 4—Estimated Family and Demographic Effects on Verbal Test Scores Between 1975 and 1990 by Racial/Ethnic Group

How Much of Test Score Changes Can Be Accounted for by Changes in Family and Demographics?

We compare our projected family/demographic effects on test scores to actual trends in NAEP test scores over similar time periods and for similar age groups to see how much of the actual changes might plausibly be attributed to changes in family/demographic characteristics. We first look at the trends in NAEP test scores. Figures 5 and 6 show NAEP score differences by racial/ethnic group between 1975 and 1990 for 13- and 17-year-old students. The results show gains in actual scores on both mathematics and verbal/reading for 13- and 17-year-old students for each racial/ethnic group.

The NAEP gains for black and Hispanic students are significantly larger than for non-Hispanic white students on each test for both 13- and 17-year-old students. Gains are from 0.45 to 0.65 of a standard
deviation for black students, 0.15 to 0.50 of a standard deviation for Hispanics, and around 0.00 to 0.15 for non-Hispanic whites. The gains for black and Hispanic students have significantly narrowed the gap between them and non-Hispanic white students—although a large gap remains. Figures 7 (mathematics) and 8 (verbal/reading) show these reductions in the test score gap among racial/ethnic groups.

We subtracted the predicted change in test scores (due to family/demographic effects) from the actual change in NAEP scores to compute a residual effect. Figures 9 (mathematics) and 10 (verbal/reading) show these residuals. The data for mathematics show no residual gain for non-Hispanic white students, indicating that their gains in test scores could be accounted for entirely by family effects. But there are large positive residuals for Hispanics and black students, suggesting that changing family characteristics alone cannot explain the large gains made by these students. In fact, changing family characteristics account for only about a third of the total gain.
For verbal/reading scores, the data generally indicate smaller residual gains than for mathematics, but still show substantial black and Hispanic residual gains not accounted for by family effects. The verbal/reading data also show that non-Hispanic white students have a small negative residual for both age groups, indicating that their NAEP gains were not as large as would be expected from family changes.

**DISCUSSION OF RESULTS AND FUTURE RESEARCH**

**The American Family**

We have used a single measure—test scores—to view the effects on youth of changes in the family. While other measures of children's development may show different results, this measure provides no evidence of a deteriorating family environment for youth who were 14–18 in 1990 compared to youth who were 14–18 in 1970/1975.
Since family influence starts early in a child's life and probably has a cumulative impact, the analysis essentially compares families from the late 1950s and 1960s to the 1970s and 1980s. Although dramatic changes have occurred in the characteristics of American families in this period—some positive, some negative—attention has focused almost exclusively on the changes perceived to be detrimental to children. The families in 1990 have more highly educated parents with fewer children and similar levels of family income compared to the families in 1970/1975. These characteristics are strongly related to student achievement and are the primary reason for predicted test score gains due to changes in family/demographic characteristics.

Although our results show that average real family income changed little over this period, this average masks two significant changes. Family income has been maintained for many two-parent families only by having two wage earners, and family income declined significantly for many children in going from a two-parent to a single-parent family. However, other characteristics of a family can be changed when a transition is made from a two-parent to a single-
parent family or from a nonworking to a working mother. For instance, these decisions can change attained educational levels, family size, and the timing of births. So the interpretation of the effects of increased numbers of working mothers and single-parent families must take account of their indirect effects on other variables as well as direct effects.

Our analysis accounts for these indirect effects of increased numbers of working mothers and single-parent families as well as a direct effect. Our results indicate that the direct effects on achievement are very small from increased numbers of working mothers and single-parent families.

The lack of a direct effect from the structure of a single-parent family is a little surprising. However, this may partly be explained by previous, but unmeasured conditions which existed for children currently in single-parent families when they were in their original two-parent family. For instance, sustained marital conflict, often involving children, can occur before divorce. Sustained conflict
within a family can significantly affect children’s development, and children who live in single-parent families resulting from divorce are probably much more likely to have been exposed to this detrimental environment in their original family. Thus, for these children, the transition to a single-parent family may not have direct negative consequences, and, in some cases, might even create a better developmental environment. Thus, the lack of a direct effect on achievement—one once other family differences are accounted for—from being in a single-parent family cannot be extrapolated to imply that children who live in a nonconfictual or positive two-parent environment would do as well in a single-family environment.

Our analysis focuses on changes in families with children aged 14–18 between 1970 and 1990. Some believe that the family environment may have worsened for younger children, particularly during the last ten years. There is some evidence to suggest that actual test score gains and estimated family effects are smaller for younger age groups and for more recent time periods (1980–1990). Further research is ongoing to see if the conclusions in this report are also true for
Figure 10—Residual Difference Between NAEP and Family Effects on Verbal Test Scores for Different Racial/Ethnic Groups, 1975–1990

younger children (age 6–10) over the same time period, and whether trends and effects are changing in more recent periods.

Test Scores as Indicators of School Quality

Comparisons of simple, unadjusted test scores from one year to the next or across different schools or districts do not provide a valid indicator of the performance of the teachers, schools, or school districts unless the differences in scores are very large compared to what might be accounted for by changing demographic or family characteristics. This is rarely the case; so, any use of unadjusted test scores to judge or reward teachers or schools will inevitably misjudge which teachers and schools are performing better. Indeed, the evidence provided here hints that a stronger case could be made that teachers and schools with large numbers of minority students may have been responsible for the most significant gains in test scores over the last 20 years, while family effects—not schools—may have been responsible for gains in nonminority scores. Although more research is
needed to test these hypotheses, this evidence illustrates the possibility of dramatic changes in perspective that more detailed analyses can provide.

Likewise, simple unadjusted nationwide test scores can be misleading as a basis for judging school quality, even if the statistical sample of students taking the tests represents the U.S. student population. Since family and demographic effects can affect scores as well as schools, the test score trends need to be adjusted for such effects before preliminary judgments are made about schools.

Using the SAT scores as a “report card on American education” is even more tenuous, since the SAT does not even draw a valid statistical sample of U.S. students. The test and samples of students taking it were never designed to provide indicators of national trends in achievement, quality of schools, or a report card on American education. Moreover, the two flaws in the statistical sample—an expanding proportion and changing composition of high school students taking the test and exclusion of non-college-bound students—both bias the test scores downward. Excluding the non-college-bound students means that the SAT misses those students making the largest gains over the last 25 years.

The purpose of the SAT test is to improve the college admissions process by providing scores that are comparable across individual students. As long as comparisons are restricted to individual students, the test can provide useful information about students applying to college. However, any aggregation of test scores above the level of the individual student—by high school, school district, state, or the nation—is simply uninterpretable as a measure of student achievement trends or of quality differences among schools, school districts, or states.

An unfortunate fact is that the public perception of school quality is partly shaped by the ever-available, but flawed, SAT scores. SAT scores can strongly influence public perceptions because they are more familiar, repeated frequently, have salience to people’s lives, and often support existing opinions. Reporting of aggregated, unadjusted SAT scores for high schools, districts, states, or the nation appears not only to serve no useful public purpose, but to confuse and detract from what should be a well-informed public debate
about our families, schools, and students. Terminating the publication of unadjusted aggregated SAT scores might also give more emphasis and resources to the more statistically accurate national tests.

**Improving Data for Resource Allocation in Education and Social Policy**

Significantly better estimates of family, school, community, and social policy effects on test scores could be obtained if there were one data set that regularly gave national tests and collected associated data from students, parents, schools, and communities. Although the NELS and NLSY data include this information, they do not provide trend data over time.\(^9\) Such data could be collected through an expanded NAEP data-collection effort, with greater information about parents, students, schools, communities, and district and state educational policies. A companion NAEP sample that starts with younger children and follows them longitudinally also seems essential for better evaluation of specific interventions and modeling of the rate of learning for individual children over time. These changes would require significant restructuring of the NAEP design as well as a significant new data-collection effort entailing significant increases in costs.

However, such data collected over time could provide trend data for almost all key variables that affect test scores of school-aged children. The result would be that changes in test scores could be more reliably related to changes in families, schools, communities, children's health, and district and state educational policies. The longitudinal component starting with younger children and larger samples of at-risk children would support much better evaluation of intervention programs as well as a much better understanding of the origins of poor school performance.

---

\(^9\) Data from the Department of Education make it possible to analyze the senior samples of the National Longitudinal Study of the High School Class of 1972 (NLS-72) and the 1992 wave of the National Education Longitudinal Survey (NELS). These data sets have comparable data on test scores and family, school, and community characteristics.
Better data and research not only help families make better educational decisions, but such data can support a much better allocation of the $300 billion or more of public funds spent annually on K–12 education and on family, social, and community programs. The additional funding required to significantly expand the NAEP would pale in comparison to what could be saved through improved private and public resource-allocation decisions that the data could support.

Assessing the Potential Effects of Public Policies and Investment on Student Achievement

Explaining minority test score gains. Our analysis shows that minorities made significant gains in test scores over the last 20 years and that a large proportion of these gains was unexplained by family changes. The dramatically rising test scores of minorities have resulted in a significant closing of the achievement gap between minority and nonminority youth and less inequality in educational outcomes. While this is a significant educational accomplishment, a significant gap remains. It is important to understand what factors contributed to these gains and whether they will continue to close the gap in the future.

Hypothesized factors that might explain the residual gains must meet four criteria. First, the hypothesized cause must either be empirically linked to test scores or at least plausibly linked to having an influence on test scores. Second, the factor must have significantly changed for youth who were 14–18 in 1970/1975 versus those who were 14–18 in 1990. Third, it must have affected black and Hispanic scores significantly but had essentially no impact on non-Hispanic white scores. Fourth, it must be a factor that would not be reflected through changing family characteristics.10

Changing public policies in the area of equal educational opportunity and increased public investment in children and schools fit

---

10 To take an example, a suggested candidate for explaining the residual might be decreased viewing of television. If so, one would have to show that students who watch less television score higher, that decreased watching occurred between 1970/1975 and 1990, that the decrease was large for black and Hispanic students but not for non-Hispanic white students, and that the decreased viewing was not simply predicted by higher education level of parents (significant effect in a multivariate model).
these criteria. The policies directed toward providing equal educational opportunity certainly would have been different for 14–18-year-old youth in 1970 who went to school in the late 1950s and 1960s as opposed to similar youth in 1990 who went to school in the late 1970s and 1980s. The barriers that have been removed to equal educational opportunity for blacks primarily involve access to integrated and probably more competitive K–12 schools, and improved access to higher education. While Hispanics might have also been affected by these policies, the policies insuring that language was not a barrier to educational opportunity might also be expected to have an impact on their test scores.

Since the major implementation of these policy changes occurred in the 1970s and 1980s, they would be expected to affect those children growing up and attending school in the time period that corresponds to our sample of 14–18-year-old youth in 1990. The effects of these policy changes would also be expected to primarily affect minority scores but to have little impact on non-Hispanic white scores, and would probably not be reflected through changed family characteristics.

Increases in public investment in K–12 schools and social programs directed at families and children also meet these criteria. There were dramatic increases in real public investment in schools and children from the 1960s through the 1980s. For example, Fuchs and Reklis (1992) estimated that per-capita (real 1988 dollars) public spending on children was $1,289 in 1960 compared to $2,946 in 1988 (about a 3 percent increase per year). Part of this additional spending was specifically targeted toward minority and/or youth from lower-income families. Since a greater proportion of black and Hispanic families have lower incomes, these programs would be expected to differentially affect minority scores.

Although some of the increased spending on K–12 education was not specifically directed toward minority youth, school administrators and teachers may have allocated more resources to lower-scoring youth. Since a greater proportion of black and Hispanic students have lower scores, minority youth would again be expected to be

---

most affected if this occurred. Social programs such as Headstart and child health and nutrition programs directed at children and families also fit this pattern of differential effects for minority youth, although the effect of some of these programs may have been proxyed by family variables included in our models.

In addition to public investment and public policies, other possible explanations for minority gains include migration trends that may have placed minorities in different school districts, changing motivation of minority students, and perhaps differential returns to education for minority versus nonminority youth.

One promising area of research is the development of improved statistical models incorporating the effects of multiple risks on children. The child-development and more clinically oriented literature repeatedly focuses on the deleterious effects of multiple risk factors on children. The hypothesis implies that test scores might fall dramatically (nonlinearly or exponentially) when children are under conditions of multiple risk. Basically, this means that the combined effects on test scores of low family income and low levels of parental education may be much larger than the additive, independent effects of each factor. Some students, then, may experience a slippery slope where achievement falls drastically with every new risk they encounter. But the reverse side of the argument is that for such children, as risks decline, student achievement should go up dramatically as well. The latter, if true, might help explain rapid gains in test scores for lower-achieving children. We are exploring the implications of the multiple-risk hypothesis.

**Explaining non-Hispanic white results.** The lack of a residual (mathematics) or a small negative residual (verbal/reading) for non-Hispanic whites needs to be explored further. One interpretation is that family effects for nonminorities may be incorrectly captured by a linear model, because marginal differences in income, family size, and parental education affect higher-scoring youth less than lower-scoring youth.\(^{12}\) If the family effects are smaller in reality, then there

\(^{12}\)We have run fully interactive models with squared terms and used these for estimating family effects. While the more complex models generally make only very small changes (less than 0.02 of a standard deviation) in the estimated size of the family effects, it should be noted that the largest change is for non-Hispanic white students.
would be larger positive residuals for non-Hispanic white youth, who tend to have higher test scores, on average.

A second hypothesis is that the public spending and changed public policies simply did not benefit non-Hispanic whites for several reasons. One is that the most effective policies and programs may have been directed primarily at minorities (desegregation, affirmative action, bilingual programs, etc.). In addition, programs targeted toward all lower-scoring students may have been less effective for non-Hispanic white students because lower-scoring non-Hispanic white students are more often located in rural areas. If rural areas have not received a proportionate share of resources and attention, students there may not have benefited to the same extent as students in urban areas. It may also just be more difficult to help lower-scoring rural youth due to their dispersion and the lack of economies of scale.

A third explanation is that while lower-scoring non-Hispanic white youth benefited from public investment, higher-scoring youth lost ground for several reasons. For instance, some have suggested a weakening of the curriculum for higher-achieving youth.\(^\text{13}\) Also, there may have been an implicit tradeoff in producing the large gains for minority or lower-scoring youth. Successfully addressing the problems of lower-scoring youth may have resulted in less emphasis and resources for higher-scoring students. These are all important issues that can be explored through future research.

**The Quality and Productivity of Schools**

This study does not support the view that schools of the 1970s and 1980s have deteriorated in significant ways with respect to the schools of the 1950s and 1960s in their instruction in mathematics and verbal/reading skills. Moreover, it suggests that schools have made significant progress in decreasing inequalities between minority and nonminority students. There have been several significant changes in schools in this time period, including school consolida-

---

tion, large real increases in per-student expenditures, integration of students, changing curriculum, smaller class sizes, and a more experienced and better paid teaching force. Some of these changes may have played a key role in boosting the scores of youth, particularly minority youth.

However, the results are not as positive from the perspective of educational productivity. The concept of educational productivity—similar to economic productivity—measures whether learning (output) per unit of resources (input) is rising or falling. Learning can change either because more years of education are achieved or because more learning occurs per year of education. Learning also has a dimension of breadth and depth. Students today may learn a wider array of subjects than students of the past, or they might learn each subject in more depth. There are obvious tradeoffs among the various components of educational productivity. For example, resources can be devoted to keeping youth in school longer, to teaching a broader array of subjects, or to focus more resources on teaching fewer subjects in depth.

While educational productivity has increased with regard to the increased student completion rates over the last 20 years, much less is known about the tradeoff between depth and breadth. For instance, students in the 1950s and 1960s did not spend time learning computer skills, and time tradeoffs may occur between learning new subjects and acquiring less in-depth knowledge of older subjects. Unfortunately, there are no good overall measures of the breadth of student knowledge. Moreover, we do not know the precise level of changing resources devoted to instruction in these areas. It is possible that these added resources were used primarily to add to the breadth of subjects, not their depth.

If we assume a constant level of resources and time devoted to subjects, our analysis suggests that there appear to be no dramatic gains in the educational productivity of schools as measured by the mathematics and verbal/reading test score trends. However, these issues need to be explored further before a full assessment can be made regarding the productivity of schools. Such an assessment would need to take into account changes in curriculum (breadth and depth), time and resources devoted to instruction, and school climate, and how these contribute to educational productivity over time.
FINAL COMMENTS

As discussed above, this study has highlighted several questions that need to be answered through future research. Developing more complex statistical models that more accurately model children's development will further clarify the importance of the family, school, and community contexts and their contribution to childhood outcomes. From our discussion, it is clear that we need to be cautious when using averages across all students to gauge changes in test scores. An average tends to obscure the fact that some groups of students may have markedly different results and that conditions for some may have worsened, lowering their achievement scores. Our results should not be interpreted to mean that conditions have improved for every student, family, or school—only that there has been a positive change when averaging across all 14-18-year-old students over the last 20 years.