DOCUMENT RESUME

ED 419 293 EA 029 055

AUTHOR	Grissmer, David W.
TITLE	Education Productivity.
INSTITUTION	National Education Knowledge Industry Association, Washington, DC.; Council for Educational Development and Research, Washington, DC.
SPONS AGENCY	Department of Education, Washington, DC.
PUB DATE	1998-00-00
NOTE	47p.
CONTRACT	RJ96006001; RJ96006101; RJ96006301; RJ96006701; RJ96006801
AVAILABLE FROM	NEKIA Communications, 1200 19th Street, N.W., Suite 300, Washington, DC 20036; phone: 202-429-5101.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Accountability; Elementary Secondary Education; Evaluation Needs; Information Needs; *Measurement Objectives; Outcomes of Education; *Productivity; *School Effectiveness

ABSTRACT

The perception exists that the massive infusion of resources into public education has done nothing to bolster student achievement scores and that American students' scores on international assessments rank far below the scores of students from other countries. These perceptions are examined in this booklet. The focus is on education productivity and the various parameters of this evaluative concept. It offers a definition of education productivity and asks whether it is a useful concept to employ. The text claims that productivity research yields the most important information for policymakers in education and that these individuals need to know how to use limited resources in the most cost-effective manner, as well as what additional outputs would be achieved with additional resources. It has been shown in communicating with corporate America, taxpayers, and others that the education community must develop and monitor credible and understandable school productivity measures and that different policies and programs be compared on the basis of cost-effectiveness. Furthermore, it is claimed, austere budgets make research that is oriented toward productivity and cost-effectiveness even more important since it allows a higher level of debate regarding public education. (Contains 62 references.) (RJM)

Education Productivity

by David W. Grissmer

Education Productivity

by David W. Grissmer

EdTalk is published by
NEKIA Communications
1200 19th Street, NW • Suite 300 • Washington, DC 20036
(202) 429-5101

Roy Forbes, Acting Executive Director
Üllik Rouk, Editor
Jessica N. Kennedy, Layout Artist

This project has been funded at least in part with federal funds from the U.S. Department of Education under contract numbers RJ96006001, RJ96006101, RJ96006301, RJ96006701, RJ96006801, and RJ96006901. The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Education Productivity
is published in conjunction with the following education R&D institutions:

Appalachia Educational Laboratory
P.O. Box 1348 • Charleston, West Virginia 25325 • 304/347-0400

Mid-continent Regional Educational Laboratory
2550 South Parker Road, Suite 500 • Aurora, Colorado 80014 • 303/337-0990

North Central Regional Educational Laboratory
1900 Spring Road, Suite 300 • Oak Brook, Illinois 60521 • 630/571-4700

Pacific Resources for Education and Learning
828 Fort Street Mall, Suite 500 • Honolulu, Hawaii 96813 • 808/533-6000

SouthEastern Regional Vision for Education
P.O. Box 5367 • Greensboro, North Carolina 27412 • 910/334-3211

Southwest Educational Development Laboratory
211 East Seventh Street • Austin, Texas 78701 • 512/476-6861

WestEd
730 Harrison Street • San Francisco, California 94107 • 415/565-3000

Preface

Are our students' achievement scores going up or down? How do our students fare when matched against students from other countries? What does the research say about the productivity of our schools as opposed to productivity in other sectors of our society? What are we getting for our money?

This *EdTalk* publication examines these questions through a long-view education productivity lens. Some of the answers may be surprising.

The Council for Educational Development and Research is made up of some of the nation's foremost institutions in the education knowledge industry, including regional educational laboratories and national education research centers. These institutions are helping educators turn findings from education research and development into successful classroom practices and are synthesizing knowledge from research and practice into useful information for education policymakers.

By informing a variety of audiences about nationally significant topics in education, the Council's *EdTalk* publication series complements these institutions' work. Our purpose in this particular publication is to spark discussion about the accuracy of the perceptions that the public and even the research community have about the effectiveness of our education investments.

The choices that we make in placing our resources as we move into the 21st century will determine whether our education productivity slows or grows. This, in turn, will govern the quality of life for all Americans well into that century and perhaps even beyond it.

One way to help ensure that we make the best choices in education is to first consider the data.

Education Productivity

by David W. Grissmer

Critics accuse the nation's public schools of using their resources inefficiently and of failing to improve "productivity." The case is best presented by Eric Hanushek (1994a; 1996a). This perception of schools contrasts sharply with how the public views other sectors of our society, where it points proudly to dramatic gains in the quality and productivity of our farms, manufacturers, and the computer industry. In an age when technological advances are pushing productivity like never before, our schools seem to lag far behind.

Are schools truly inefficient in their use of resources? Have the investments we made in the Great Society years added up to nothing? There is probably no more important set of questions in public education than those related to school productivity. School productivity research could tell us whether additional resources make a difference in student achievement, whether allocating additional resources to some programs is more effective than allocating them to others, and which types of students benefit from more or different resource allocations.

Answering these questions is critical to determining whether public education needs more resources, needs to allocate its resources differently, or needs to fundamentally restructure in order to use its resources more effectively.

A driving force behind this paper was the perceptions that the public and some in the research community have about K-12 education. One of these perceptions is that the massive infusion of resources

A driving force behind this paper was the perception that the public and some in the research community have about K-12 education.

has done nothing to stop student achievement scores (as measured by SAT scores) from falling, particularly among minority students. Consequently, money makes no difference. A second perception is that American students' scores on international assessments rank far below the scores of students from other countries. Consequently, Japanese schools, for instance, are far more productive than American schools. A third perception about schools is that private schools achieve higher test scores than public schools and that they do it with fewer resources. Consequently, private schools are more productive than public schools.

If all these perceptions were correct, it certainly would seem that public education is simply not able to improve its productivity and utilize resources well. A solid case might be made for restructuring school governance so that resources could be used more effectively. However, under close scrutiny, some of the conclusions about public schools are much more favorable than public perceptions would imply. For others, there are clear reasons, seeped deep down in our culture and beliefs, for why American students appear not to do as well as students in some other countries.

Some of the conclusions about public schools are much more favorable than public perceptions would imply.

We begin by reviewing the concept of productivity as it is defined in the economic sense and as it is applied to our private sector firms and industries. We then discuss the strengths and limitations of applying the concept to education. We look at its application in four contexts: schools versus private sector industries, schools in the mid-1960s to the early 1990s, American versus Japanese schools, and public versus private schools. Next, we address the importance of incorporating the concept of productivity into education research and provide some examples where it would be useful. Finally we turn to the implications for research if productivity is going to be more than another passing education fad.

Throughout this paper we focus on "education" productivity rather than "school" productivity. The former concept encompasses all sources of learning and support, including the most important component of productivity in learning — the family.

Defining Productivity

Productivity involves not only "doing better," but doing better with equal or fewer resources. If resources increase and outcomes improve, it does not necessarily follow that productivity improves. For instance, doubling the number of cars a factory manufactures by doubling factory capacity keeps productivity at the same level as it was before. Productivity increases when we increase output while holding inputs constant. Or another way to increase productivity is to reduce inputs (downsize), yet still manage to keep outputs stable. Measuring productivity always involves measuring some outcome or output quantity per quantity of input. For instance, we measure labor productivity in our national economy by the value of the goods produced per hour of labor input.

Productivity involves not only "doing better," but doing better with equal or fewer resources.

Perhaps the best, long-term example of productivity gains is in the farming sector. Farm productivity has increased markedly during this century, whether measured by output per hour of labor or output per unit of arable land. This gain in productivity is commonly attributed to advances in farm technology, better seeds, weed and pest control, improved management, and increased economy of scale from larger farms. One result of this increased productivity is that the number of farmers and farm laborers has declined as a proportion of the workforce. Since each unit of labor can produce so much more, fewer are needed.

But, as this example illustrates, productivity can be a double-edged sword. On the one hand, producing more output per unit of labor input means that more goods and services can be available to society and the standard of living higher. In fact, economists believe that higher standards of living in the long term can only result from gains in productivity. That is partly why we collect an extensive amount of economic data to measure labor productivity, and why our economy provides strong incentives to increase productivity.

On the other hand, increased productivity in the absence of a stronger demand for goods often requires fewer workers. Thus, employment can fall in those very industries that make the most rapid productivity gains. U.S. industries may have recently experienced this phenomenon when increased worker productivity, some of it due to computers, enabled them to downsize.[1] Ideally, if the economy is robust, new jobs created in different industries, aided by worker retraining programs, can absorb this displaced labor.

Private Sector Productivity: Implications for Education

In the long term, growing productivity is a product of our capitalistic economy. The rate of productivity growth, however, can vary markedly over time and among different sectors of the economy. Placing school productivity in the context of private sector productivity helps to establish reasonable expectations for productivity growth and explain productivity trends and differences among schools and school districts.

The expectation has been that education reforms, innovation, and new technology would cause school productivity to rise. The labor productivity of U.S. manufacturing workers rose fairly steadily from 1949 to 1973 at a compound growth rate of about 1.8 percent per year (*Monthly Labor Review*, 1995). Had this trend continued, output per worker would double about every 39 years. However, private sector productivity experienced historically small increases after 1973. From 1973 to 1992, it grew at only a compound rate of 0.8 percent. At this rate, worker output would double only every 89 years. Moreover, there appears to have been negative productivity growth between 1973-1979.

During the period when schools were most criticized for not improving productivity, the U.S. as a whole was experiencing abnormally low levels of productivity growth.

No consensus exists on reasons for this slowdown in the growth of productivity in manufacturing (Wolff, 1996). Some believe that the energy crisis and associated higher oil prices and inflation were part of the cause. Others say that slackening innovation and inadequate investment in new capital, partly caused by low savings and higher interest rates, are to blame. Still others see inadequate workforce skills, partly due to poor education, as a component of the slowdown. Whatever the cause, it is important to remember for our later discussion that during the period when schools were most criticized for not improving productivity, the U.S. as a whole was experiencing abnormally low levels of productivity growth. It is possible that the same factors that retarded growth in the private sector also retarded growth in schools, especially when we take into account that schools have characteristics similar to those private sector firms and industries that traditionally have the slowest productivity growth.

It is also important to realize that the rate of productivity growth differs greatly among industries. Slow productivity growth usually masks an underlying dynamic where some industries have little or no productivity growth while others have very rapid growth. For instance, industries experiencing the most rapid productivity growth between 1973 and 1992 (growth rates were over 2.5 percent per year) were manufacturing computers, electronics, and electrical equipment. Manufacturers of transportation equipment, furniture, and metal products experienced the slowest growth, less than 0.4 percent per year.

Higher productivity occurs in industries that are "capital" intensive as opposed to "labor" intensive.

One fundamental premise about productivity articulated by economist William Baumol is that higher productivity occurs in industries that are "capital" intensive as opposed to "labor" intensive. Underlying this hypothesis is the simple fact that it is usually easier to improve the productivity of machines than it is of people. Activities that are labor intensive are essentially those where we have not found machines to replace people. A common example is haircuts — a very labor intensive activity. We do not expect the productivity of barbers to increase much over time since it takes about as long today to give a haircut as it did 20 years ago. In those industries and occupations, like barbers, we expect slower or little productivity growth.

Industries that produce goods rather than services, however, continually find ways to increase productivity, first by replacing people with productive machines and then by building even more productive machines. According to this theory, we would not expect education to be among industries with rapid productivity growth because it is very labor intensive — most education expenditures go to people rather than capital.

One implication of the labor intensive nature of education is that the regular Cost of Living Index (CPI) is an inadequate instrument with which to adjust education expenditures over time to real dollars (Rothstein and Miles, 1995). Later, we will discuss the implication of this for school productivity.

Education Productivity Research: Two Critical Questions

Almost all the research discussion has been about how best to spend the money, when how much money should the nation, state, or school district invest in education may be more important.

Measuring productivity in education is more difficult than it is in the private sector, both theoretically and practically. Perhaps the most important difference is that in education, productivity analysis must answer two tough questions, while in the private sector it must address only one. Private sector economic productivity analysis only needs to address how to best produce a given level of output. The "optimal" level of output for a firm or industry is presumably specified by the market demand for its product. If the competitive market works, the level of demand for the particular type of goods made by a firm or industry will be "optimal." In education, no market specifies how much education output is enough. Consequently, we must ask not only how much should we spend on education but also how we should spend it.

By asking these questions, we attest to the fact that we can err by spending too much or too little money on education, as well as by spending what we have on the wrong programs or policies. A productive education system not only allocates its resources well, but also spends the right overall level of resources to achieve the "optimal" level of education output. So far, however, almost all the research discussion has been about how best to spend the money — the cost-effective mix of education programs and policies — when how much money should the nation, state, or school district invest in education may be more important.

Research on this latter question requires estimating the social and economic costs and benefits that accrue from education investments and estimating the potential rate of investment return if spending levels were higher. Such estimates are common in health care where, for instance, the social and economic costs of smoking are arrayed against the cost of prevention programs. The cost of prevention is treated as an investment that generates a long-term return in lower health care costs and more earnings from a longer and healthier work life. Some studies of this type were conducted on the investment in Head Start (Berrueta-Clement et al., 1984). Other work estimated the impact of education quality on economic productivity (Bishop, 1989) and economic returns to education (Card and Krueger, 1992; 1994). With credible estimates, we could generate a rate of investment return and compare it to alternative uses of public and private funds. If investing in education provided higher rates of return, then higher funding levels would prove a more productive investment for society.

With credible estimates, we could generate a rate of investment return and compare it to alternative uses of public and private funds.

Barriers to Applying the Concept of Productivity to K-12 Education

Measuring productivity in private sector industries, though not as complicated as in education, is still a complex procedure. First, it requires measuring both the inputs and the outputs of a process. Second, it must link only those inputs that produce the outputs. Third, to accurately measure output per single input, all other inputs must be held constant. Finally, the output in question must reflect only what a particular firm adds to the product — not the complete value of the product. This "value added" concept is essential to evaluating the productivity of particular firms or industries.

For instance, the price of a car includes all the costs associated with its manufacture, from the extraction of the ore to make steel to the final assembly. Hundreds of firms produce the parts that eventually make up that price. Nonetheless, we can identify how productive a particular firm is in doing its specific part — or to put it another way, to measure the value it adds to the cost of the car — by subtracting the cost of individual inputs from the price of the output. This tells us how much value that firm adds.

Education is directed toward the cognitive, emotional, and social development of children and youth and, as such, defies easy measurement.

The private sector is geared to producing a variety of goods which are automatically valued in dollars. Schools, on the other hand, produce a variety of outputs that are not amenable to aggregation in a common measure. Education is directed toward the cognitive, emotional, and social development of children and youth and, as such, defies easy measurement. Most schools do not test children's emotional and social development, or measure skills like cooperation, communication, and teamwork. In the cognitive area alone, where the most effort to measure learning progress has been made, knowledge spans multiple subjects, each involving several layers of complexity. The best way to capture fairly what a child has learned even in a single subject is still subject to controversy. Comprehensive tests that measure both depth and breadth of knowledge are rare — but critical if we are to develop improved measures of productivity in education.

A second big problem is separating the contribution that schools make to education from that of families, communities, and other sources of education. This distinction is crucial since differences in children's learning, as measured by achievement scores, are mainly due to factors outside schools — primarily the family. Separating these contributions from those of schools to derive "value added" measures may prove too analytically complex even if data were available to make the separations. A major problem here is that if we use test scores as measures of output, we need to collect — at minimum — important family variables along with the scores to impute fair and accurate productivity measures to schools themselves. This escalates our data requirement and may run into issues of privacy.

It may be possible to impute family effects from alternate sources of data like the U.S. Census (Grissmer et al., 1994). However, we need a lot more research before we can develop fair and accurate value added measures for state school systems and district school systems. Small sample sizes make measures at the school or class level problematical especially given the fact that student migration into and out of schools during a year might impact class or school score averages as much as other effects.

Another barrier to measuring school productivity effectively is the inability of current education budgeting and accounting systems to link inputs with expected outputs. As a result, there is virtually no state or district where finance reporting conventions permit the aggregation of expenditure data in ways that would shed light on the purpose of the expenditure (Guthrie, 1996). Education research in this area has traditionally studied inputs and outputs separately. The assessment research has focused only on measuring outputs while school finance research has analyzed only inputs. Failing to specify the purpose of various types of expenditures makes productivity analysis nearly impossible and may be the reason for education production function studies showing such variance.

Education research in this area has traditionally studied inputs and outputs separately.

For instance, if we do not distinguish between resources devoted to "socially desirable objectives" and academic objectives, measures of the effects of resources on achievement are going to be biased downward. And there is solid evidence (Rothstein and Miles, 1995; Lankford and Wyckoff, 1996) that a very sizable fraction of new education resources from 1970 to 1990 were spent on socially desirable objectives like special education. These expenditures would not be expected to boost regular students' achievement and thus should not be categorized as inputs to raise overall achievement scores.

The simplest type of school productivity measurement that would take all these factors into account involves annually collecting the following data from a sample of schools:

• focusing on a single measurement of output — an achievement score in a single subject;

• measuring the test scores at the beginning and end of the school year to determine a value added during the school year;

• collecting data throughout the year from teachers, students, and families involving various subject-specific, time-on-tasks measures (classroom time, homework, parental time input); and

• collecting other input data like teacher and family characteristics, class size, supplemental teacher aides, characteristics of other students and family, school expenditures, and community social capital.

Some states are moving toward the collection of data that would form the basis for productivity measurement. However, even with the appropriate data, there are significant analytical issues in turning the data into value added measures. The most difficult of these is properly controlling for the effects of family inputs. Traditionally, family effects have been "controlled for" by a SES variable or linear family characteristic variables. However, family effects are much more complex than those captured by simple linear variables. Previous work in this area may contain a common set of specification problems.

An alternative approach is to measure "education productivity," a measure that includes all sources of learning or support for learning, and not try to separate out family and community effects.

The Debate About Education Productivity

The perception of a productivity crisis in American education stems from several sources.

Americans have come to expect steady improvements in productivity as a way of life — and indeed, that is the way it has been for over a century. But now, there is the sense that declines in education productivity may be threatening the standard of living that these earlier improvements generated.

The perception of a productivity crisis in American education stems from several sources. First, there is a common perception that K-12 education received very large increases in real resources from the mid-1960s through the 1990s while achievement scores declined. Figure 1 shows the data underlying this perception. The graph on the left shows a dramatic rise in overall per pupil spending. In 1994, the average per pupil expenditure was about $6300, almost double what it was in 1967. The data is adjusted by the Consumer Price Index (CPI) to convert to real dollars. The graph on the right shows dropping Scholastic Aptitude Test (SAT) mathematics and

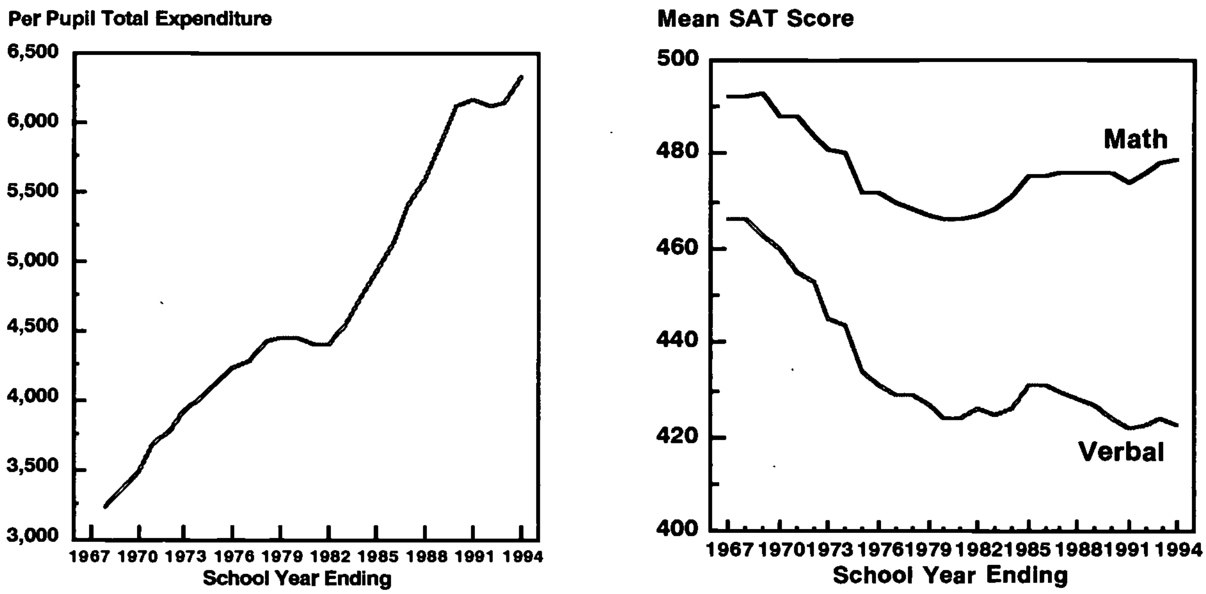


Figure 1 — Trends in Per Pupil Expenditure and Mean SAT Scores

verbal scores over the same years. Even though mathematics SAT scores have rebounded recently, average scores remain markedly below those of 25 years earlier.

Because education spending has increased and SAT scores have decreased, the public perceives a negative return on its education investment. And, in fact, if Figure 1 does accurately depict resources (inputs) and results (outputs), it would, indeed, imply a virtual collapse of education productivity in this country. However, as we shall see very shortly, these data do not provide accurate measures of either inputs or outputs.

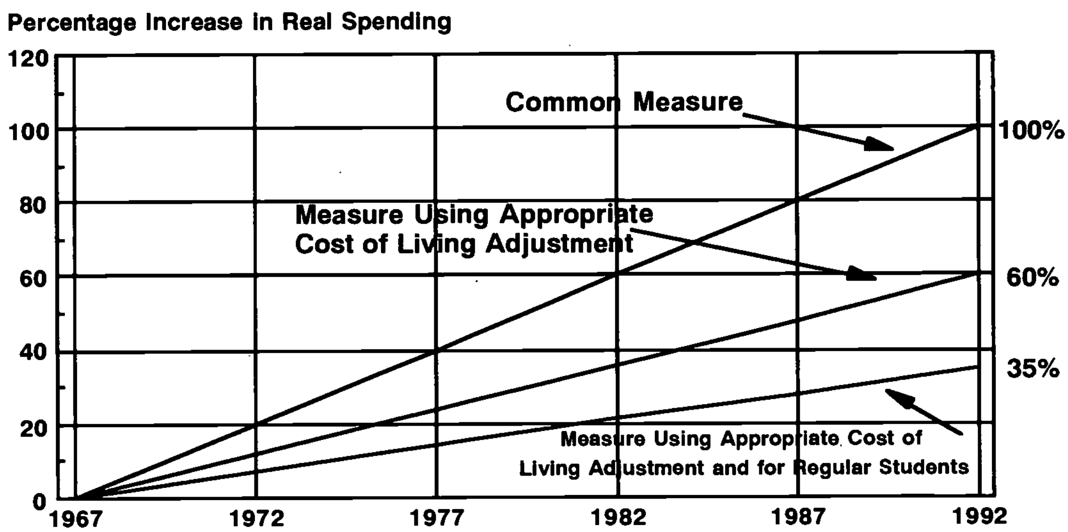
A second set of evidence that has influenced perceptions about education productivity is Hanushek's (1989) review of empirical evidence from over 300 studies on the relationship between resources and student achievement. This review concludes that, in the end, "money does not matter" (Hanushek, 1994; 1996a; 1996b). The same hypothesis is the theme of another recent book that focuses on the use of economic applications in education policy (Burtless, 1996). Although Hanushek doesn't directly say so, the results imply that productivity could be increased by cutting school resources just to the point where lower resources actually produce lower levels of outputs. Hanushek prefers the much milder conclusion that

Because education spending has increased and SAT scores have decreased, the public perceives a negative return on its education investment.

The spending increases are significantly lower than the 100 percent increase in per pupil spending that is so frequently cited.

schools do not need more money but need to reallocate present funds. Let's examine both the evidence in Figure 1 and Hanushek's review of empirical studies to see if the facts really support the perception that education productivity collapsed despite increased funds.

We turn first to trends in per pupil expenditures. Figure 2 shows the percent increase in per pupil spending from 1967 to 1992. The top line depicts the common measure typically cited for school spending increases: Between 1967 and 1992 school spending increased by 100 percent in real terms. However, research is beginning to re-estimate the size of the increase in education spending and what it has bought. For example, Rothstein and Miles (1995) found that when adjusting for cost-of-living changes more specific to the education sector, total per pupil expenditures increased by 60 percent between 1967 and 1992 (the middle line in Figure 2) — still a substantial increase. When the researchers made additional adjustments to estimate the increase in spending for regular students (i.e., not special education students), per pupil expenditures increased by approximately 35 percent over the 25-year period, as shown in the bottom line in Figure 2. These adjustments more accurately describe the education spending increases over time. The



Source: *Where's the Money Gone?*, Economics Policy Institute

Figure 2 — Percentage Increase in Real Per Pupil Expenditure

spending increases are significantly lower than the 100 percent increase in per pupil spending that is so frequently cited. Rothstein and Miles' study also shows that the largest expenditure gains for non-special education students were directed toward lower income or minority students. Consequently, these groups of students would be expected to show the largest score gains.

Although SAT scores most often drive public opinion about national test score trends, SATs are seriously flawed as indicators of changing achievement among American students. First, the SAT sample is not representative of U.S. students. The number of high school students taking SATs has increased from approximately 30 to 43 percent, introducing a downward bias. Also, a constantly changing proportion and composition of students take SATs, again introducing a downward bias in scores over time. And, finally, from our perspective, a more serious flaw is that the SAT sample excludes students not going to college. However, the largest learning gains in the late 1960s to early 1990s have occurred among students who have generally been considered low achievers, students who are less likely to go to college and/or to take the SAT. Thus, SAT scores exclude the very population making the largest gains (Berliner and Biddle, 1995; Powell and Steelman, 1996).

SATs are seriously flawed as indicators of changing achievement among American students ... [they] exclude the very population making the largest gains.

A far better measure of student achievement than the SAT is the National Assessment of Educational Progress (NAEP) — an assessment designed specifically to monitor trends in U.S. students' achievement (Koretz, 1986). NAEP consists of a set of standardized tests in core subjects. Administered by the Department of Education since the early 1970s, it is taken by nationally representative samples of students aged 9, 13, and 17. The test items used for comparing achievement have remained stable over time and so yield more accurate data than tests where content has changed.

How do NAEP test score trends compare to those of the SAT? Figure 3 shows conflicting results for verbal scores, with SAT scores declining roughly 8 percentile points and NAEP scores gaining nearly 4 percentile points. Trends in mathematics are in closer agreement but still differ by about 3.5 percentile points.

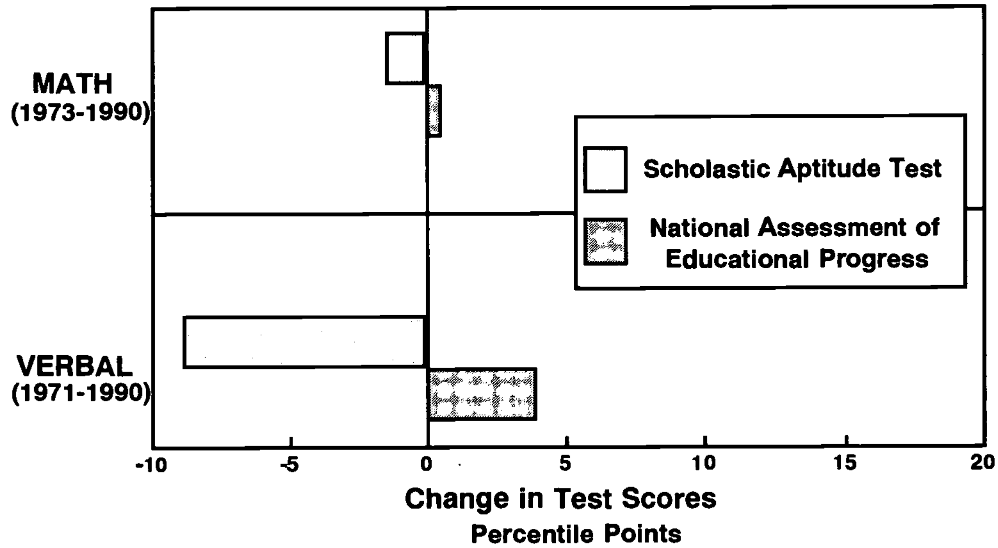


Figure 3 — Trends in Student Achievement: SAT and NAEP

While NAEP math and reading scores both increased, increases varied significantly for different racial/ethnic groups (Figure 4). Black and Hispanic students posted the greatest improvements. For instance, among 17-year-olds, non-Hispanic white students gained about 4 percentile points, black students gained about 23 percentile points, and Hispanic students bettered their scores by 7.5 percentile points. Similar patterns appear for other age groups and for reading scores, although the magnitude of the gains differs, especially for 9-year-olds.

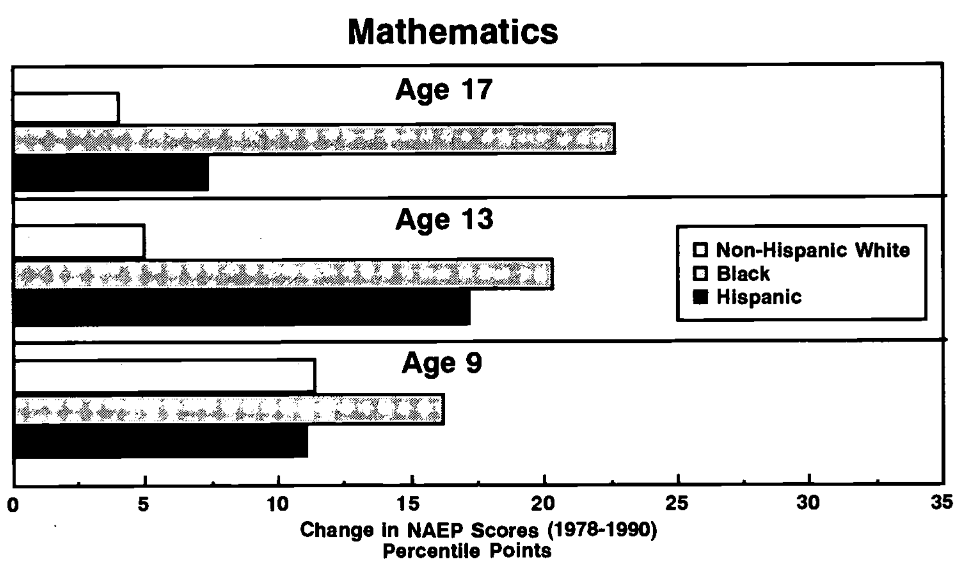


Figure 4 — Change in NAEP Mathematics Scores By Racial/Ethnic Group

The significant gains made by black and Hispanic students relative to non-Hispanic white students has helped to narrow the math score gap between minorities and nonminorities. Nonetheless, a substantial difference remains (Figure 5).

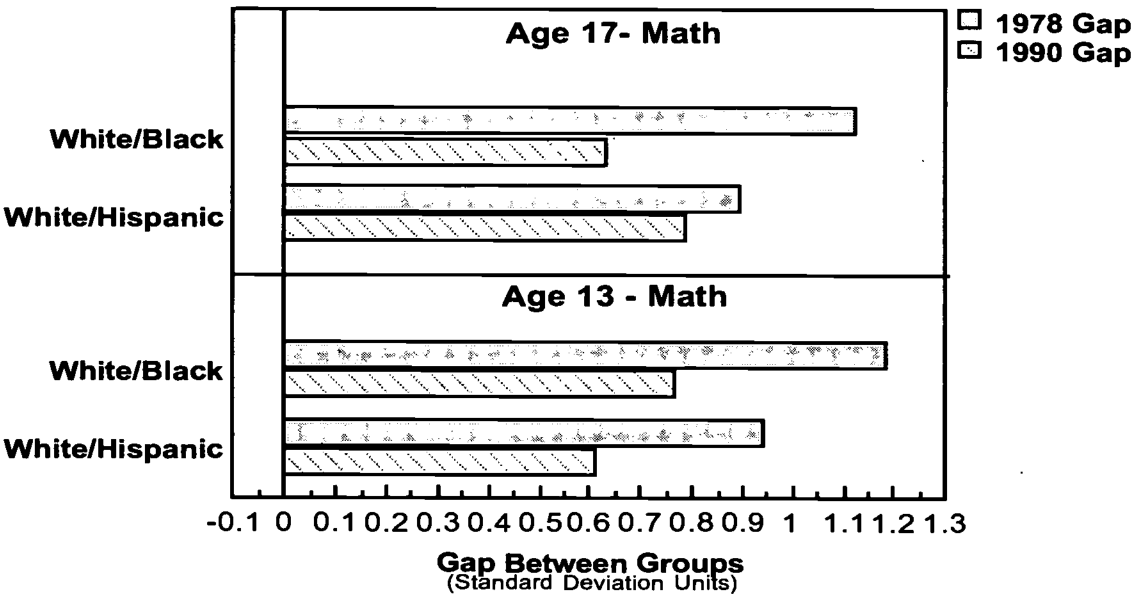


Figure 5 — Gap in NAEP Mathematics Scores in 1978 and 1990 By Racial/Ethnic Group

These trends in NAEP raise critical questions. First, what accounts for such increases in test scores if it is true that the condition of the American family has deteriorated over the past several decades? Second, what accounts for the significant achievement gains by minority students? Perhaps the most viable hypothesis is that public investments in families and schools and/or equal education opportunity policies have yielded important payoffs. If so, identifying which programs have worked and their relative cost-effectiveness, especially for students placed at risk of education failure, stands out as a critical topic for further research.

To determine if social and education investments might be responsible for increased scores, we need to identify and separate out effects that changes in the family might be expected to have on NAEP scores. However, because NAEP does not collect important family variables, alternate data sources (Grissmer et al., 1994) are being used to estimate family effects.

21

The Changing Family and Student Achievement

Balancing those family characteristics that are more supportive of student achievement against those that are less supportive, we predicted that the overall impact of family changes on student achievement is positive.

Sorting out family characteristics that contribute to student achievement trends is a complex exercise because we must consider several factors simultaneously. Our analysis consists of four steps: (1) estimating the magnitude of changes in family/demographic characteristics over the past 20 years for representative samples of youth; (2) estimating the relationships between student achievement scores and family/demographic characteristics with cross-sectional regression models; (3) using these relationships to predict test scores for three national samples of youth between the 1970s and 1990; (4) comparing these predicted changes to actual changes in student test scores over the same period.

These steps both capture the net effect of each of the family and demographic variables on student achievement scores and incorporate the degree of change in these variables between two generations of families. By making separate estimates for black, Hispanic, and non-Hispanic white families, it is possible to determine whether the effects of family changes are different for minority and nonminority families.

The methodology used here consisted of three steps: (1) developing equations relating student achievement to family and demographic characteristics using The 1980 National Longitudinal Survey of Youth (NLSY) and the 1988 National Education Longitudinal Study (NELS), both of which are large nationally representative datasets; (2) utilizing these equations to predict test scores for each student in a national sample of children (from the Current Population Surveys) in 1970, 1975, and 1990 using their family and demographic characteristics; and (3) comparing the mean differences in these predicted test scores (estimates of the effect of changing family and demographic characteristics) to actual scores from the National Assessment of Educational Progress (NAEP). This procedure provides an estimate of how much changing family and demographic changes contributed to actual changes in test scores. Residual changes in test scores provide an estimate of the "value added" from factors not related to family and demography.

Based on available data and prior research, we analyzed several family characteristics and their effect on test scores, including parents' education levels, family size, family income, the age of the mother at the birth of the child, mother's employment status, and whether the child lives with a single parent (see also Grissmer et al., 1994). Figure 6 illustrates how we compute the total net effect of changes in selected family characteristics on predicted test scores. To determine the influence of these characteristics requires examining the direction and magnitude of change over time (column 2) and the net impact of each family characteristics on test scores (column 3). Considering both the amount of change and the amount of net influence on test scores reveals the combined effect on test score trends. Thus, the final column shows the magnitude and direction of each family characteristic's relationship to student scores in the 20-year period.

Family Factor	Amount of Change (1970-90)	Amount of Net Influence on Test Scores	Combined Effect on Test Scores
Parental Education	Large	Large	Large ↑
Family Size	Large	Medium	Medium ↑
Family Income	None	Medium	None
Mother's Age at Birth	Small	Medium	Small ↓
Working Mother	Large	None	None
Single Parent	Large	Indirect	Indirect
			Net Family Impact ↑

Figure 6 — Estimating the Net Effect of Changing Family Factors

Balancing those family characteristics that are more supportive of student achievement against those that are less supportive, we predicted that the overall impact of family changes on student achievement is positive.

In Table 1 we first examine the magnitude of the changes in family characteristics for different racial/ethnic groups. The dramatic increase in parent education levels and the marked decline in family

Table 1
Selected Family Characteristics of 14 to 18-Year-Olds, 1975-1990*

	Black	Hispanic	White
	Percent Change (1975-1990)		
Mother's Education (%)			
Less Than High School	-53	-12	-44
College Degree	154	61	76
Father's Education (%)			
Less Than High School	-58	-11	-54
College Degree	221	-12	42
Number of Children			
1-2	111	38	42
4 or More	-71	-43	-66
Median Family Income ($)	-2	-21	-1

*** 1975 was the first year Hispanic students and families were identified in the data.**

size are important in explaining test score trends among all groups. Declines in family size coupled with level average family income (in real terms) between the mid-1970s and 1990 means that family income per child actually increased during this time period.

Black families experienced more favorable changes than non-Hispanic white and especially Hispanic families. The percentage of black parents without a high school diploma has decreased substantially, while the percentage with a college degree has increased by 150 to 220 percent. Another factor contributing to improved test scores by black students' test scores is the significant decline in the size of black families.

Changes in Hispanic families were less positive than in other racial/ethnic groups. Family income levels among Hispanics declined in real terms by about 12 percent. Changes in parents' education levels and family size were less dramatic.

In addition to changes in family characteristics, we examined the relative importance of each family factor on student test scores over time.

Figure 7 shows unadjusted group differences in achievement among children with different family/demographic characteristics on the NELS and NLSY survey data. Large differences occur according to parents' education attainment. For example, we find that children of college graduates score about 35-40 percentile points higher on mathematics tests than children whose parents who did not graduate from high school. Large differences also occur between different racial/ethnic groups, with black students scoring 30 percentile points lower in mathematics than non-Hispanic white students, and Hispanic students scoring approximately 22 percentile points below non-Hispanic white students.

Large differences [in student achievement] occur according to parents' education attainment.

Income, mother's age at child's birth, family size, and single versus two parent family status all appear related to student achievement. Differences of approximately 5-10 percentile points emerge among the different groups.

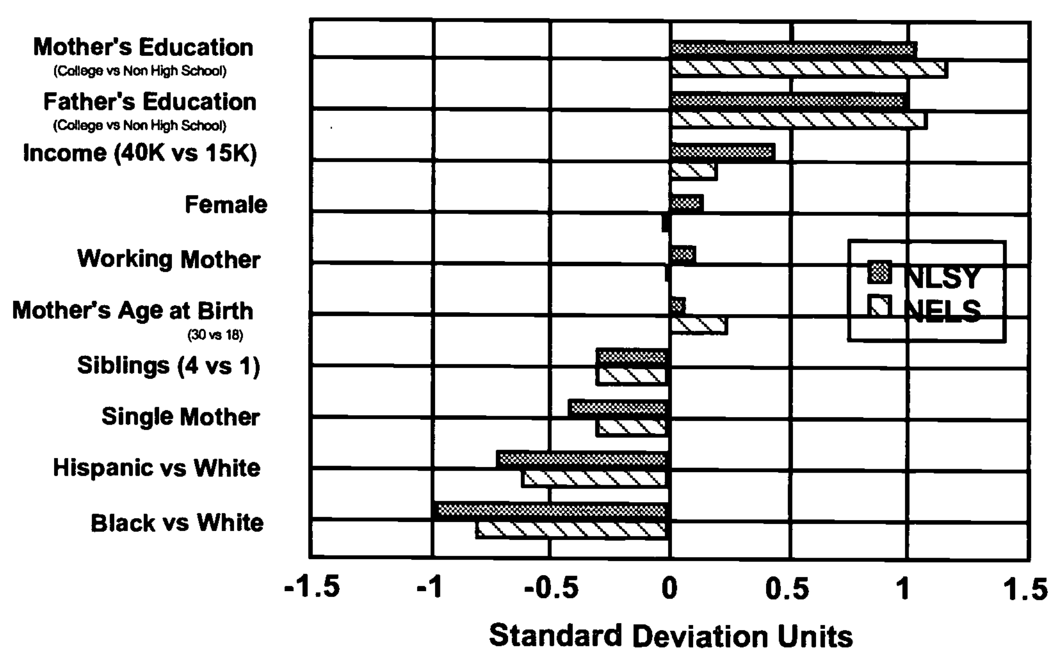


Figure 7 – Simple Differences in Mean Mathematics Test Score for Selected Groups, NLSY and NELS

There was essentially no difference between the achievement of children with working mothers versus those with mothers who did not work outside the home.

Children who come from households with a family income of $40,000 score 5 percentile points higher in mathematics than children from families with incomes of $15,000. Children of older mothers (30 or more years old at time of birth) score 5 percentile points higher in mathematics compared to children of teen mothers (18 years old at time of birth). Children with a greater number of siblings do worse on tests by about 10 percentile points.

Children in households with single mothers score about 10 percentile points below those in two-parent households. There was essentially no difference between the achievement of children with working mothers versus those whose mothers did not work outside the home.

Relying on simple, unadjusted group differences may lead to erroneous inferences, however. The problem is one of confounding factors. For instance, lower test scores among children in single parent households may reflect other, more important, family conditions such as family income. What we really need to determine is the net relationship between each family characteristic and student achievement. To show the differences between unadjusted and adjusted effects, Figure 8 compares the unadjusted test score differences for selected groups of children with the test score differences that would exist if the children had otherwise similar characteristics but differed in this one variable alone (i.e., adjusted or net relationships).

We find that overall, the net effects tend to be much smaller — less than one-half of the gross effect. This is not surprising given that the net effect controls for the effect of other variables while the gross effect includes them. The pattern in terms of overall importance on test scores remains much the same, however. Where earlier we saw test score differences of over 35 percentile points between children whose parents did not have a high school diploma and those whose parents were college graduates, we now predict differences of about 18 percentile points, which, while large, are not as large as the unadjusted effects.

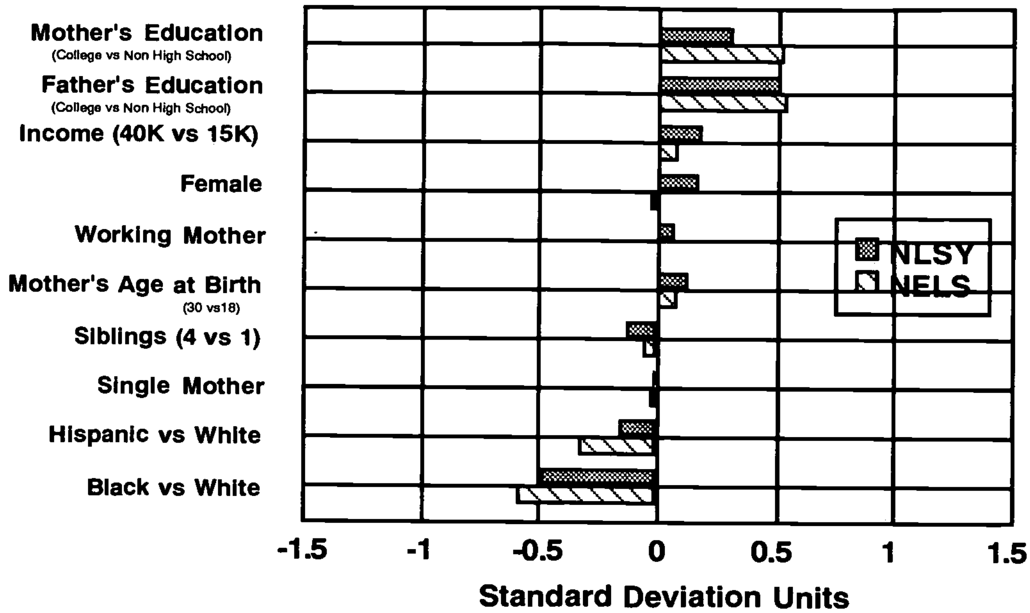


Figure 8 – Net Differences in Mean Mathematics Test Scores for Selected Groups, NLSY and NELS

Especially noteworthy is the fact that the effect on achievement of being brought up in a female-headed household is essentially zero, very different from the large difference that appeared earlier. Apparently, a lot of the unadjusted difference is due to income, low maternal education levels, and other factors that frequently characterize single parent families rather than family structure itself. The differences by race/ethnicity are still quite large, but considerably smaller than the unadjusted effect.

Using unadjusted effects almost always overstates the effect of a variable and in some cases implies an effect that disappears under controlled conditions. Thus, advocating policies based on the use of unadjusted effects can be very misleading.

The effect on achievement of being brought up in a female-headed household is essentially zero.

Based on the magnitude and direction of family changes and the relative influence of different family characteristics on student achievement, students in 1990 would be predicted to score higher, not lower, on tests than youth in families in 1975 (Figure 9). Between 1975 and 1990, non-Hispanic white students and black students would have gained 6 percentile points. For the same time period, the predicted test score gains for Hispanic students were about 4 percentile points less.

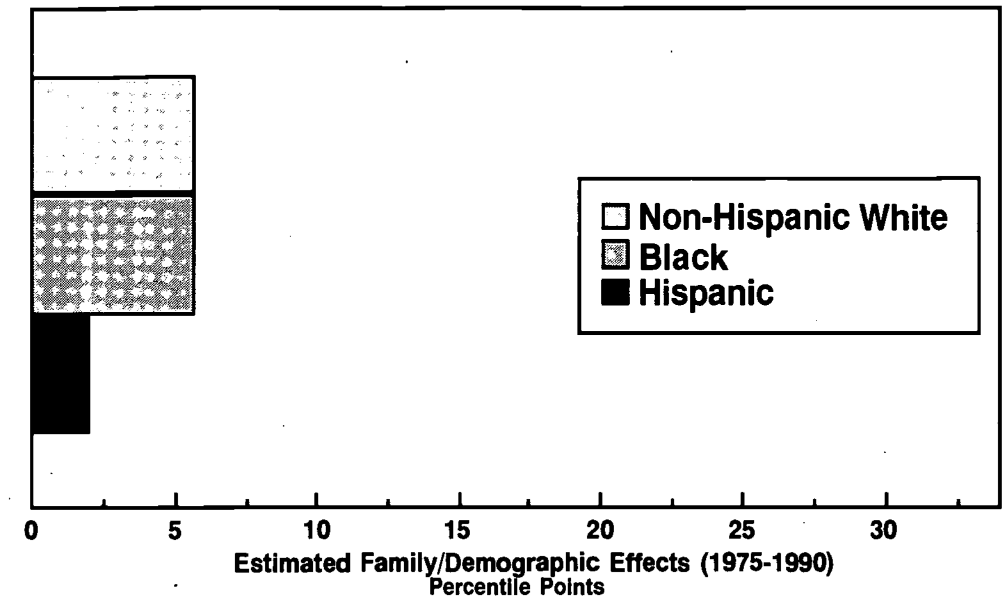


Figure 9 — Change in Predicted Mean Mathematics Test Score for Different Racial/Ethnic Groups, 1975-1990

Family characteristics alone cannot explain the large gains these students made.

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. Figure 10 shows the residuals for mathematics for the period from 1978 to 1990. (The 1978 test was the first NAEP mathematics test to identify Hispanics.) There is no residual gain for non-Hispanic white students. This indicates that family effects might entirely account for their gains in test scores. However, there are large positive residuals for Hispanics and black students, suggesting that changing family characteristics alone cannot explain the large gains these students made. In fact, changing family characteristics account for only approximately one-third of their total gain. We need to look at other factors to help explain the other two-thirds.

In summary, our analysis of national test score trends highlights improvements for various age groups between 1970 and 1990. (See Grissmer, 1994, for results of all NAEP mathematics and verbal tests from 1971 to 1990.) All racial/ethnic groups have contributed to positive test score trends. While the test score gains of non-Hispanic white students have been modest, the gains of minority students have been substantial, with black students experiencing the largest gains.

These positive test score trends are due in part to improved family conditions such as parents' higher education attainment, smaller families, and more income per child. Because family changes explain only about one-third of the test score gains for minorities, it is likely that minority youth have benefited from other factors (for example, the social and education investment and policies aimed at minority and low-income families) during the late 1960s through early 1990s.

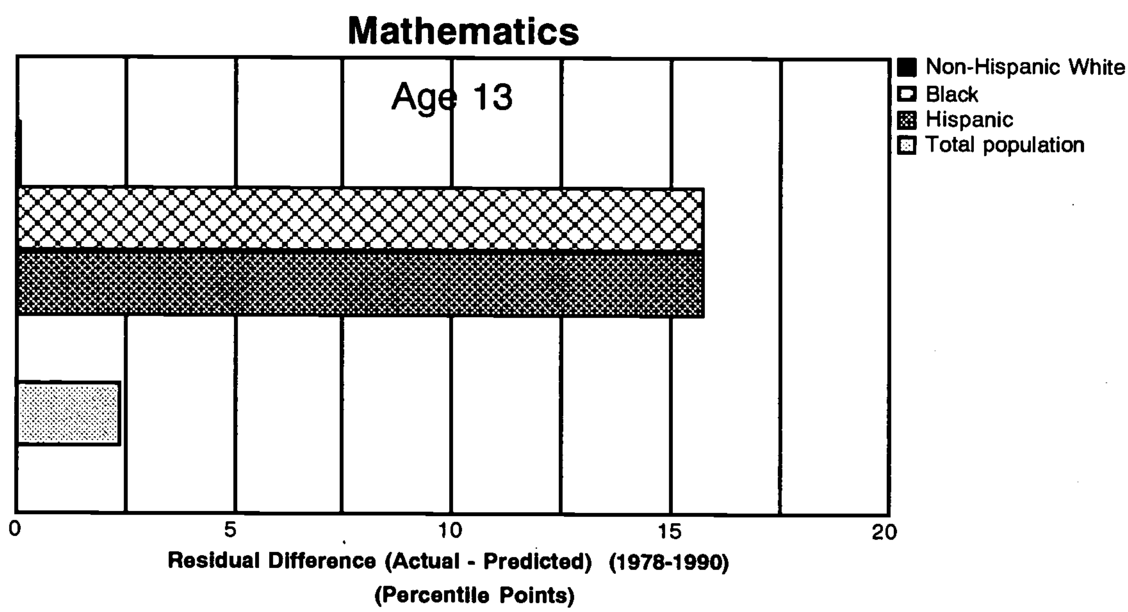


Figure 10 — Residual Differences Between NAEP and Family Effects on Mathematics Test Scores for Different Racial/Ethnic Groups

Evidence about what students have achieved and what schools have cost during this period challenges the widely held view that student performance has deteriorated despite massive infusions of resources. These two studies point to the possibility that student performance actually rose in the 1970 and 1990 period while resource increases were much smaller than commonly perceived. Minority students may have actually improved their performance. In other words, precisely the students who received the largest increase in resources may have achieved the greatest gains in test scores.

Precisely the students who received the largest increase in resources may have achieved the greatest gains in test scores.

Other recent work at the state level shows positive relationships between resources and achievement (Ferguson, 1991; Ferguson and

Ladd, 1996). Evidence from an experiment also shows significant student gains when resources are spent to lower class size in the early grades (Mosteller, 1995). Minority students have twice the gain from lower class size as do nonminority students.

The Pattern of Minority Score Gains

An interesting question is whether black students' gains are cohort specific or period specific (Koretz, 1986). A cohort effect would show that gains first occurred for 9-year-olds, then four years later for 13-year-olds, and in another four years for 17-year-olds. It is developmentally based in that it assumes that a score at any age is a cumulative result of environmental conditions from birth. So if, for instance, better prenatal care resulted in higher birth weight, we would expect to see this effect for all age groups born after the implementation of such a prenatal program. A period specific effect would occur for all age groups in the same year.

Analyses of NAEP data using a cohort perspective offer several useful insights. Figures 11 and 12 present NAEP reading and math scores for black students according to the year they entered school.

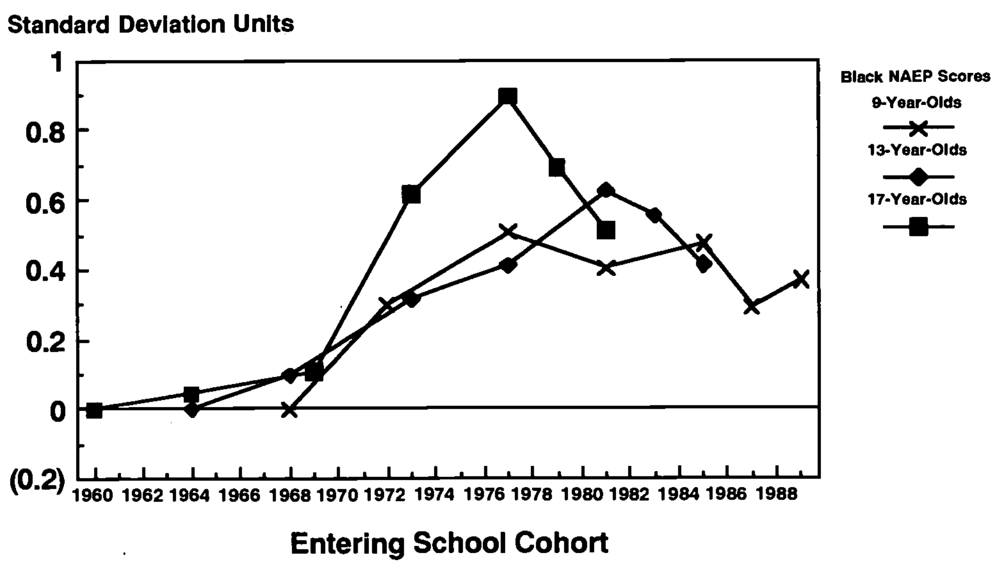


Figure 11 – Change in NAEP Reading Scores by Entering School Cohort and Age: Blacks

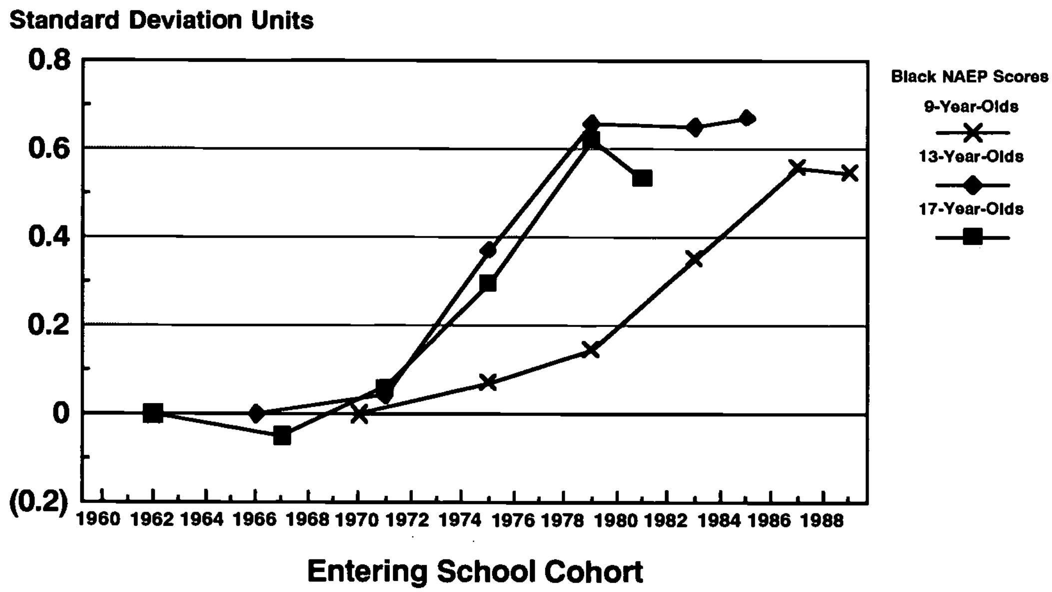


Figure 12 – Change in NAEP Math Scores by Entering School Cohort and Age: Blacks

These data allow us to compare scores of a single cohort at ages 9, 13, and 17. The data display a pattern of fairly stable scores for black students who entered school before 1968, rapid gains for black students who entered between 1968 and 1978, and little or no gain, and even some decline, for students who entered school in 1980 and after. For instance, the cohort that entered school in 1976 scored about 0.4 standard deviations higher in reading at ages 9 and 13 than did students who entered school in 1968, and almost 0.8 standard deviations higher at age 17. Thus, students who entered school in 1976 show greater reading gains at each age that they took the test than do students who entered school in 1968. The cohorts that entered school after 1980 show no additional gains in scores, and some decline. However, their sustained gains are in the 0.4 standard deviation range.

Students who entered school in 1976 show greater reading gains at each age that they took the test than do students who entered school in 1968.

NAEP math data also display a pattern of rapidly rising scores for students who entered school in 1975 and 1979 compared to those who entered school in 1968. Increases occurred at all ages within these cohorts — although they were much more pronounced for 13 and 17-year-old students. The test scores of 13 and 17-year-olds also stabilized after 1980, showing no additional gains. However, the scores of 9-year-olds alone continued to increase after 1980 — an exception to a cohort effect. Students who entered school in

The data focus attention on the question of what differences did students who entered school in 1970 and before experience compared to those who entered school in 1980 and later.

1983 and 1987 showed increased scores at age 9, but there were no more increases after that.

The data focus attention on the question of what differences did students who entered school in 1970 and before experience compared to those who entered school in 1980 and later. Certainly a hypothesis that needs examination is whether the gains that later cohorts made corresponded to the implementation of new practices such as compensatory education for disadvantaged students, reductions in class size, or equal opportunity programs passed in the 1960s under civil rights, "Great Society," or "War on Poverty" legislation.

The Ongoing Productivity Debate

How does one make sense of this mixed evidence? On the one hand, we have rising test scores among minority students in the 1970 to 1990 period, when class size nationally was substantially reduced, equal opportunity and compensatory programs were implemented, and higher levels of social spending occurred. On the other hand, we have the weight of over 300 empirical, nonexperimental studies pointing to no average effects from higher spending.

The current approach to analyzing the effects of resources and policies on achievement is to simply weigh the evidence from hundreds of studies (Hanushek, 1989; 1994a; 1996a; 1996b) equally. As new findings emerge, they shift the conclusion slightly toward one side or the other. However, there are large differences in the quality and assumptions of these studies. Some consideration has to be given to weighting this evidence according to appropriate quality standards.

One approach is to use quality or similarity criteria to group studies according to what resources are being tested, how dependent and independent variables are defined, how models are specified, and the characteristics of the student population being tested. This would provide some indication of whether differences in results can be due to these kinds of differences. It is certainly possible, and maybe probable, that poor quality data and poor specifications of variables and models are responsible for significant biases that run through most of the empirical studies. A recent article (Hanushek, 1996b) begins to examine the role of aggregation in biasing outcomes.

Measuring the relationship between resources and achievement requires specifying the appropriate education resource, family and student outcome variables, and statistical models. However, nearly all national data collection efforts have lacked key elements with which to make sound measurements. NAEP data have been collected since 1971, but NAEP does not collect information on family variables that influence achievement or resource data from schools. The Department of Education's longitudinal data collections (High School and Beyond and NELS) lack good resource measures. The NLSY data set tested with the Armed Forces Vocational Battery in 1980 also lacks school resource measures.

Very low-scoring students often come from multiple risk family situations, and that each additional risk produces a compounding effect.

Moreover, there is a common set of specification problems that runs through these previous studies. One is the failure to include multiple risk in family specifications. Recent work (Grissmer et al., 1995) indicates that very low-scoring students often come from multiple risk family situations, and that each additional risk produces a compounding effect. Virtually all previous statistical models have neglected this effect. Failure to incorporate multiple risk in statistical models has the potential to significantly bias measurements of resources and achievement. It is particularly important to accurately portray the effects that families have on low-scoring students in resource equations since studies show that these students are the ones expected to make the largest gains.

Second, almost no previous empirical, nonexperimental study makes allowances for contextual family effects. This means, for instance, that the specifications for the model assume that an additional dollar of income has the same achievement effect on a child of a non-high school educated single teen mother with three children as on an only child of a two college-educated parent family. Alternately, linear family models assume that it makes no difference whether a single parent is college educated or lacks a high school degree, or whether a working mother has one or five children. Virtually all linear models of education achievement lack the precision to capture important family contextual effects — which could potentially be an important source of bias.

Third, almost no models recognize that achievement is the cumulative result of all previous schooling and family environments.[2] For instance, a test score in eighth grade depends on class size not just in eighth grade, but in all the grades before it — and may depend more on class size in earlier grades. Most models use single grade contemporaneous measures of resources rather than measures reflecting the experience over a student's entire school career. This can introduce still another significant source of bias.

A test score in eighth grade depends on class size not just in eighth grade, but in all the grades before it.

Fourth, Rothstein and Miles' research shows that the most common measure of resources — per pupil expenditures — can be seriously flawed. When used in time series measurement and adjusted by CPI inflation factors, this measure overstates the real increase in resources, which — other things being equal — biases the effects of resources downward. In addition, previous measurements rarely included variables for how resources are utilized programatically. If we do not distinguish between resources devoted to "socially desirable objectives" and academic objectives, then the effects of resources on achievement are also going to be biased downward. And there is solid evidence from Rothstein and Miles' work and other studies (Lankford and Wyckoff, 1996) that a very sizable portion of increased resources from 1970 to 1990 went to socially desirable objectives like special education. These expenditures would not be expected to boost achievement of regular students.

Finally, and perhaps the main reason that education production function studies have shown different results lies in the way resource variables are defined and collected. Finance reporting conventions do not permit school district or state expenditure data to be aggregated in ways that shed light on the efficacy of school strategies.

Virtually no previous nonexperimental study meets these objections. Therefore, it is not unreasonable to conclude that the accumulated body of evidence may be sufficiently flawed so as not to be able to measure the effect of resources on student test scores with any precision. Experimental studies, if well executed, would not be subject to any of these problems. The single experimental study undertaken (Mostellar, 1995) shows significant test score effects for smaller class size. While one would like more than one experiment, it is possible that this measurement should "outweigh" all nonexperimental investigations.

Fortunately, the situation is changing and significantly better empirical studies are becoming possible. Many states are annually testing students at multiple grades. State tests have several advantages over national test data in measuring the relationship between resources and achievement. First, the variance in NAEP scores and resources across districts in a state is much greater than across states. Second, samples at the individual school and district are much larger than interstate samples. Third, tests for multiple grades within each school allow better tracking of resources over a student's career. Fourth, it is easier to track resource levels and where resources are spent at a state level than nationally. This is because the different tracking systems that states use make interstate comparability problematical. Finally, many states have excellent data on teachers, which can be matched to students and schools and tracked over time. This means that teacher characteristics can be much better identified in equations.

What we need is a study that explains why results differ so widely, and to illustrate with a single, coehensive set of data that different variable and model specifications can produce different results, and that certain variable and model specifications can better explain the variance in achievement scores, especially for lower-scoring and minority students. With newly emerging state data sets, this will be possible.

Finance reporting conventions do not permit school district or state expenditure data to be aggregated in ways that shed light on the efficacy of school strategies.

American versus Japanese versus Taiwanese Schools: Which Are More Productive?

If we take a productivity approach and measure test score per unit of time spent on the subject, it is not clear that American schools were lagging in productivity.

The comparisons that are frequently made between Japanese and American schools illustrate the importance of taking a productivity perspective in education. While there is strong evidence that Japenese students score higher than do American students on mathematics and reading tests at a given grade, it is not clear that Japanese schools are more productive than American schools. Most standard international tests prior to 1996 failed to collect data that would explain why these differences exist; that is, whether they relate to better school productivity, higher school resources, or factors in the family environment, for instance. The research (Stevenson and Stigler, 1992; Stevenson et al., 1993) that has collected comparative data on families, classrooms, and test scores begins to explain these differences and places the results into more of a productivity framework. Stevenson and Stigler's 1992 study estimated value added in test scores in a sample of Japanese, Taiwanese, and American schools by testing first and fifth-grade students at the beginning and end of the year. Thus, it estimated an output measure that depended only on the inputs during a single year.

The results showed that the biggest difference in Japanese, Taiwanese, and American students' education was in "time on task" variables. Both Japanese and Taiwanese students had significantly more homework, more time in the classroom, and more outside tutoring than American students. All of these activities contributed to more time on task. However, if we take a productivity approach and measure test score per unit of time spent on the subject, it is not clear that American schools were lagging in productivity.

How much time children spend in school is clearly not a variable that schools themselves control. It reflects a deeper set of cultural beliefs about children and curriculum. The results of the study may indicate that if Americans were willing to extend the school year, make the school day longer, give students more homework, and place more emphasis in the classroom on mathematics and reading, then test score gaps might decline. One possible conclusion to be drawn from all of this is that the productivity of American schools is partly a question of the cultural commitment to having children spend more time on learning.

Differences in the attitudes of Japanese and American parents clearly illustrate the power of such a commitment. This may reflect cultural factors as well as resource utilization factors. Japanese parents believe that hard work (time on task) is the primary determinant of achievement, while American parents believe that innate ability is the primary determinant (Stevenson and Stigler, 1992; Stevenson et al., 1993). Japanese parents also have much higher expectations for their children and schools and express higher levels of dissatisfaction with their children's test scores — even though their scores are much higher than American students' scores. It is clear that school productivity is a joint function of school and family traits, and that measurements of school productivity must take into account different family resources, expectations, and levels of commitment.

The productivity of American schools is partly a question of the cultural commitment to having children spend more time on learning.

The international comparisons also raise other interesting productivity issues. One has to do with the allocation of resources. Is hiring more teachers and reducing class size better than raising teachers' salaries and having large classrooms? An intense debate about these issues is taking place in American schools. The Japanese have opted for the latter option while Americans the former. Class size in Japan is almost twice that in American schools — so fewer teachers are needed — but teachers are more highly paid than in America. Japanese teachers also have significantly less contact hours per day with students and more time to prepare lectures. Japanese children spend more time during the day in recreational activities and with instructional aides than do American children. One can hypothesize from this that adequate time for teachers to prepare lessons may yield better use of learning time in the classroom. American teachers tend to have no time during the day for preparation and must prepare their lessons in the evening, essentially working many hours "overtime."

A second issue for productivity that this research raises has to do with the way in which Japanese schools manage their much larger classes. Japanese schools spend significant time in the early grades organizing and teaching children appropriate behavior and using groups to self-police discipline. Their greater emphasis on recreation and breaks during the day may also help children to be more focused and less disruptive during lessons.

The newly released Third International Mathematics and Science Study (TIMSS) for the middle years (Beaton et al., 1996a; 1996b) will allow much better analysis of these questions. Initial results present only basic tabulations of results by countries. These are arrayed by single variables describing demographic characteristics, family characteristics, teacher characteristics, and school system and school characteristics. It will require much further analysis to determine the implications of those results for the productivity of the U.S. education system. However, it should be noted that the results place U.S. students near the middle in the ranking of countries. The TIMSS evidence on inputs to achieve these rankings is somewhat mixed, but certainly does not support the hypothesis that the productivity of U.S. schools is significantly out of line with schools in other countries.

Japanese schools spend significant time in the early grades organizing and teaching children appropriate behavior and using groups to self-police discipline.

The data on inputs shows that students in the U.S. spend less time than students in the average country on homework, but spend significantly more time on sports, jobs at home, and being with friends. Class size in the U.S. is near the average for all countries, as is the time spent each week on either science or mathematics. However, two measures of school expenditures show the U.S. spending to be above average.

Since class size, a key determinant of expenditures, is about average, it is unclear whether these higher expenditure measures in the U.S. reflect differences in teacher salary levels, higher costs for non-academic objectives (special education, etc.), higher costs for administration, transportation, or other expenses, or simply problems in conversions made to a single currency. Without analyzing the expenditures further, it is difficult to judge whether real additional dollars in the U.S. are devoted to academic objectives. Overall, it appears that student time inputs are much less than average, teacher time inputs are near average, and financial inputs are above average. However, more analysis will be required to accurately describe the productivity of U.S. schools using the data.

Education Productivity: A Useful Concept?

It is important to place the debate about schools in a productivity framework for several reasons. First, productivity research yields the most important information for policymakers in education. Virtually all major education policy decisions involve both inputs and outputs. Policymakers need to know how to use limited resources the most cost-effectively and what additional outputs would be achieved with additional resources. These questions are also uppermost in the minds of corporate executives and business persons who become involved in education. Much of their frustration arises from the inability of the research and data to show what works and at what cost.

Policymakers need to know how to use limited resources the most cost-effectively and what additional outputs would be achieved with additional resources.

Second, communicating with corporate America, taxpayers, and those in the legislature and executive branches of states and localities requires that the education community develop and monitor credible and understandable school productivity measures and that different policies and programs be compared on the basis of cost-effectiveness. Taxpayers — the group whose opinion about education matters most — largely work in the private sector and face issues of productivity regularly. They are keenly aware of the implications of failing to boost productivity for themselves and their employers. A great deal of cynicism and resistance exists among corporate American and taxpayers when it comes to funding education. Many of these perceptions may be wrong, but it will take good productivity research to change it.

A third reason for examining schools through a productivity lens is that austere budgets make research oriented toward productivity and cost-effectiveness critically important. Future resources for education will likely be squeezed by rising enrollments, the fervor of tax-cutting, and conflicting demands for the use of state and local funds by criminal justice systems, social welfare programs, and infrastructure needs (Odden, 1995). Making the wrong resource allocation decisions costs more — other things equal — in terms of outputs at lower budget levels than at higher budget levels.[3] For instance, it is still possible to improve achievement even if budgets do not rise. To do so, however, we need to find ways to reallocate current budgets so that they provide more resources for programs that are more cost-effective and fewer resources for programs that are less cost-effective.

Research on productivity and the cost-effectiveness of programs and policies can play a key role in restoring trust between educators and policymakers, and between the research community and the American people, who fund education.

Finally, research on productivity and cost-effectiveness can better inform the debate about public education. High-quality research that achieves consensus can narrow the range of viable policy options to be debated and can separate issues which can be decided empirically from those that are purely ideological. Research has failed to answer too many policy questions, leading to interminable debates and demagoguery, and lurching from one new, untested policy or program to another. The string of reform failures over the last 25 or maybe even 100 years (Tyack and Cuban, 1996) has engendered deep cynicism about education among teachers, principals, school superintendents, policymakers, and taxpayers. This situation will not improve until better quality research leads to tested and proven programs and policies that are both effective and efficient. Research on productivity and the cost-effectiveness of programs and policies can play a key role in restoring trust between educators and policymakers, and between the research community and the American people, who fund education.

Americans, however, spend very little on research in education. Nationally, research and development accounts for between 2 and 3 percent of our gross domestic product. In health and transportation, research and development expenditures run between 2 and 3 percent of expenditures in these areas. Research and development consumes more than 15 percent of our defense expenditures. In contrast, only one-third of one percent of education expenditures in the U.S. are spent on education research and development. It is hard to see how test scores can go up if we do not develop the critical mass of high-quality education research that engineers education improvement.

Underinvestment in research and development and the problematical quality of some of the research are partly to blame for education's not having made large gains in productivity. A greater investment in high-quality research and development would create a stronger research infrastructure and, in turn, higher levels of productivity. The research infrastructure would need to include stronger interdisciplinary academic programs directed toward understanding the development of children from birth; the roles of families, peers, communities, and schools in producing higher education output; stronger and more integrated longitudinal data sets; and more exploration of the role of technology in raising education output (Wilson, 1994). Most people associate computers in the classroom with educational technology. However, the technology being developed to explore learning disabilities through brain imaging and development of "relearning" systems may prove in the long term to be the direction where technology can be most useful to education (Shaywitz, 1995).

It is hard to see how test scores can go up if we do not develop the critical mass of high-quality education research that engineers education improvement.

Productivity growth is usually ascribed to investment in new capital (new facilities, machines, automation, etc.), new technologies arising from research and development, and increased education, skills, and health of the workforce. However, it is research and development that cultivates the advances that get incorporated into capital investment. In the long run, successful research and development is probably the most important factor in productivity growth.

Notes

[1]Roach (1996) makes an interesting distinction between the causes of traditional productivity gains — better quality workforce and technology — and the recent productivity gains caused by downsizing, which he describes as a one-time occurrence. He points out that downsizing may actually hinder productivity and economic growth in the longer term, whereby traditional productivity gains usually left industries posed to take advantage of increased demand.

[2]An exception to this is that some difference models control for earlier test scores in a student's school career and attempt to measure resources between the different grades. However, family variables continue to be significant even if earlier scores are used for controls.

[3]This assumes that the marginal effect of rising expenditures becomes smaller.

42

References

Barro, S.M. (Spring, 1989). Fund distribution issues in school finance: Priorities for the next round of research. *Educational Evaluation and Policy Analysis,* 11, 2, 17.

Beaton, A.E., Martin, M., Mullis, I., Gonzalez, E., Smith, T., & Kelly, D. (1996a). *Science Achievement in the Middle Years,* Chestnut Hill, MA: Boston College TIMSS International Study Center.

Beaton, A.E., Mullis, I., Martin, M., Gonzalez, E., Kelly, D., & Smith, T. (1996b). *Mathematics Achievement in the Middle Years,* Chestnut Hill, MA: Boston College TIMSS International Study Center.

Berliner, D.C., & Biddle, B.J. (1995). *The Manufactured Crisis: Myths, Fraud, and the Attack on America's Public Schools,* Reading, MA: Addison-Wesley Publishing Company.

Berne, R., & Stiefel, L. (1995). *Student-level school resource measures.* Paper prepared for the National Center for Education Statistics, Annual Data Conference, Washington, DC.

Berrueta-Clement, J.R., Schweinhart, L.J., & Barnett, W.S. (1984). *Changed Lives: The Effects of the Perry Pre-School Program on Youths Through Age 19,* Ypsilanti, MI: High/Scope Press.

Betts, J.R. (1996). Is there a link between school inputs and earnings? Fresh scrutiny of an old literature. In G. Burtless (Ed.), *Does Money Matter?: The Effect of School Resources on Student Achievement and Adult Success,* Washington, DC: The Brookings Institution, 141-191.

Bishop, J.H. (1989). Is the test score decline responsible for the productivity growth decline? *The American Economic Review,* 79, 178-197.

Borjas, G.J. (1990). *Friends or Strangers: The Impact of Immigrants on the U.S. Economy,* New York: Basic Books, Inc.

Bronfenbrenner, U., & Ceci, S.J. (1994). Heredity, environment, and the question "How?" — A first approximation. In R. Plomin & G.E. McClern (Eds.), *Nature, Nurture, and Psychology,* Washington, DC: American Psychological Association, 313-324.

Bryk, A.S., Lee, V.E., & Holland, P.B. (1993). *Catholic Schools and the Common Good,* Cambridge: Harvard University Press.

Burtless, G., (Ed.). (1996). *Does Money Matter?: The Effect of School Resources on Student Achievement and Adult Success.* Washington, DC: The Brookings Institution.

Monthly Labor Review: Productivity Measurement and Trends (July 1995). Bureau of Labor Statistics, U.S. Department of Labor.

Caplan, N., Choy, M.H., & Whitmore, J.K. (1992). Indochinese refugee families and academic achievement. *Scientific American*, 266, 36-42.

Caplan, N., Choy, M.H., & Whitmore, J.K. (1991). *Children of the Boat People: A Study of Educational Success*, Ann Arbor, MI: The University of Michigan Press.

Card, D., & Krueger, A.B. (1994). The economic return to school quality: A partial survey. Industrial Relations Section, Princeton University. Working Paper #334.

Card, D., & Krueger, A.B. (February 1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of Political Economy*, 100, 1-40

College Entrance Examination Board. (1991). College Bound Seniors: 1991 Profile of SAT and Achievement Test Takers, Princeton: College Entrance Examination Board.

Consortium on Productivity in the Schools. (1995). *Using What We Have to Get the Schools We Need*, New York: Columbia University Teachers College.

Ferguson, R., & Ladd, H.F. (1996). Does money matter? Additional evidence from Alabama on how money matters. In H.F. Ladd (Ed.), *Performance-Based Approaches to School Reform*, Washington, DC: The Brookings Institution, 265-298.

Ferguson, R.F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation*, 28, 465-498.

Grissmer, D.W., Kirby, S.N., Berends, M., & Williamson, S. (1995). *Multiple risk effects of demographic, family, and other factors on cognitive achievement*. Invited paper presented at the 7th International Conference on Socio-Economics of the Society for the Advancement of Socio-Economics, Washington, DC.

Grissmer, D.W., Kirby, S.N., Berends, M., & Williamson, S. (1994). *Student Achievement and the Changing American Family*, Santa Monica, CA: RAND.

Guthrie, J.W. (1995). Implications for policy: What might happen in American education if it were known how money actually is spent? In L.O. Picus & J.L. Wattenbarger (Eds.), *Where Does the Money Go? Resource Allocation in Elementary and Secondary Schools*, Sixteenth Annual Yearbook of the American Education Finance Association. Thousand Oaks, CA: Corwin Press, Inc.

Hanushek, E.A. (1996a). Measuring investments in education. *Journal of Economic Perspectives*, 10, 9-30.

Hanushek, E.A., Rivkin, S.G., & Taylor, L.L. (1996b). Aggregation and the estimated effects of school resources. *The Review of Economics and Statistics*, 611-627.

Hanushek, E.A. (1994a). *Making Schools Work: Improving Performance and Controlling Costs*, Washington, DC: The Brookings Institution.

Hanushek, E.A. (1994b). Money might matter somewhere: A response to Hedges, Laine, and Greenwald. *Educational Researcher*, 23, 5-8.

Hanushek, E.A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy*, 100, 84-117.

Hanushek, E.A. (1989). The impact of differential expenditures on school performance. *Educational Researcher*, 18, 45-51

Hedges, L.V., Laine, R.D., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on outputs. *Educational Researcher*, 23, 5-14.

Herrnstein, R.J., & Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life*, New York: The Free Press.

Kirst, M.W. (1988). The internal allocation of resources within U.S. school districts. In D.H. Monk & J. Underwood (Eds.), *Microlevel School Finance: Issues and Implications for Policy*, Cambridge: Ballinger Publishing Co.

Koretz, D. (Winter 1992). What happened to test scores, and why? *Educational Measurement: Issues and Practice*, 7-11.

Koretz, D. (1991). State comparisons using NAEP: Large costs, disappointing benefits. *Educational Researcher*, 20, 19-21.

Koretz, D. (1987). Educational Achievement: Explanations and Implications of Recent Trends, Washington, DC: U.S. Congressional Budget Office.

Koretz, D. (1986). *Trends in Educational Achievement*, Washington, DC: Congressional Budget Office.

Lankford, H., & Wyckoff, J. (1996). The allocation of resources to special education and regular instruction. In H.F. Ladd (Ed.), *Holding Schools Accountable*, Washington, DC: The Brookings Institution, 221-257.

Levin, B. (1996). Improving the productivity of schools: Learners as the main resource. In *Making Schools Productive*, Washington, DC: Center for Educational Development and Research, 13-23.

Levin, H.M. (1983). *Cost-Effectiveness: A Primer*, Newbury Park, CA: Sage Publications.

Linn, R.L., & Dunbar, S.B. (1990). The nation's report card goes home: Good news and bad about trends in achievement. *Phi Delta Kappan*, 72, 127-133.

McLanahan, S., & Sandefur, G. (1994). *Growing up with a Single Parent: What Hurts, What Helps*, Cambridge: Harvard University Press.

Monk, D. (1992). Educational productivity research: An update and assessment of its role in educational finance reform. *Educational Evaluation and Policy Analysis*, 14, 307-332.

Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *The Future of Children: Long-term Outcomes of Early Childhood Programs*, Los Altos, CA: The David and Lucille Packard Foundation.

Murnane, R.J., & Levy, F. (October 1996). What General Motors can teach U.S. schools about the proper role of markets in educational reform. *Phi Delta Kappan*, 77, 109-114.

North Central Regional Educational Laboratory. (Fall 1996). Productive schools: Perspectives from research and practice. *New Leaders*, Oak Brook, IL: North Central Regional Educational Laboratory.

Odden, A. (1996). The costs of restructuring to high performance school design: The case of the New American Schools. In *Making Schools Productive*, Washington, DC: Center for Educational Development and Research, 35-40.

Odden A., & Clune, M.S. (1995). Improving educational productivity and school finance. *Educational Researcher*, 24, 6-10.

Picus, L.A. (1996). Productive schools and educational resource allocation. In *Making Schools Productive*, Washington, DC: Center for Educational Development and Research, 43-48.

Picus L.O., & Wattenbarger, J.L. (1995). Where does the money go? Resource allocation in elementary and secondary schools. *Yearbook of the American Education Finance Association*. Newbury Park, CA: Corwin Press, 1-19.

Roach, S.S. (November-December 1996). The hollow ring of the productivity revival. *Harvard Business Review*, 74, 81-89

Rothstein, R., & Miles, K.H. (1995). *Where's the Money Gone? Changes in the Level and Composition of Education Spending*, Washington, DC: Economic Policy Institute.

Shaywitz, S. (November 1996). Dyslexia. *Scientific American*, 98-104

Stevenson, H.W., Chen, C., & Lee, S.Y. (January 1993). Mathematics achievement of Chinese, Japanese, and American children: Ten years later, *Science*, 259, 53-58.

Stevenson, H.W., & Stigler, J.W. (1992). *The Learning Gap*, New York: Summit.

The Future of Children (Winter 1995). 5, Los Altos, CA: The David and Lucile Packard Foundation.

Tyack, D., & Cuban, L. (1995). *Tinkering Toward Utopia: A Century of Public School Reform.* Cambridge, MA: Harvard University Press.

Walberg, H.J. (1994). Educational productivity: Urgent needs and new remedies. *Theory Into Practice*, 3, 75-82.

Walberg, H.J., (Ed.). (1996). *Advances in Educational Productivity*, 5, Greenwich CT: JAI Press Inc.

Wilson, K.G. (1994). *Redesigning Education*, New York: Henry Holt.

Wolff, E.N. (1996). The productivity slowdown: The culprit at last? Follow-up on Hulten and Wolff. *The American Economic Review*, 86, 1239-1252.

Zill, N., & Rogers, C.C. (1988). Recent trends in the well-being of children in the United States and their implications for public policy. In A.J. Cherlin (Ed.), *The Changing American Family and Public Policy*, Washington, DC: The Urban Institute Press, 31-98.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

NOTICE

REPRODUCTION BASIS

☐ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☑ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").