WHAT MATTERS FOR STUDENT ACHIEVEMENT

Updating Coleman on the influence of families and schools

THE COLEMAN REPORT, "Equality of Educational Opportunity," is the fountainhead for those committed to evidence-based education policy. Remarkably, this 737-page tome, prepared 50 years ago by seven authors under the leadership of James S. Coleman, still gets a steady 600 Google Scholar citations per year. But since its publication, views of what the report says have diverged, and conclusions about its policy implications have differed even more sharply. It is therefore appropriate—from the Olympian vantage point a half century provides—not only to assess the Coleman findings and conclusions but also to consider how and where they have directed the policy conversation.

It must be said from the outset that the Coleman team relied on a methodology that was becoming antiquated at the time the document was prepared. Almost immediately, econometricians offered major critiques of its approach. But even with these limitations, as an education-policy research document, the report was breathtakingly innovative, the foundation for decades of ever-improving inquiry into the design and impact of the U.S. education system.

Outside the scientific research community, the Coleman Report had, if anything, an even broader impact. Reporters, columnists, and policymakers turned their understanding of results and conclusions into conventional wisdoms—simplified, bumper-sticker versions of the report's conclusions. Partly reflecting the nature of the document, not all of them agreed on which of the findings to emphasize. For example, early on, President Lyndon Johnson's administration said the report endorsed its desegregation efforts by showing that blacks benefited from an integrated educational experience while whites did not suffer from it. This message dovetailed with the administration's efforts to implement the Civil Rights Act, a topic discussed by Steven Rivkin in an accompanying essay (see "Desegregation since the Coleman Report," Spring 2016). Later, two other, more lasting conclusions attributed to the report gradually emerged: 1) families are the most important influence on student achievement, and 2) school resources don't matter. I focus on these two conclusions.

The greater significance of the Coleman Report—what makes it a foundational document for education policy research—lies not in any of these interpretations or conclusions, however. More importantly, it fundamentally altered the lens through which analysts, policymakers, and the public at large view and assess schools. Before Coleman, a good school was defined by its

by ERIC A. HANUSHEK

"inputs"—per-pupil expenditure, school size, comprehensiveness of the curriculum, volumes per student in the library, science lab facilities, use of tracking, and similar indicators of the resources allocated for the students' education. After Coleman, the measures of a good school shifted to its "outputs" or "outcomes"—the amount its students know, the gains in learning they experience each year, the years of further education graduates pursue, and their long-term employment and earnings opportunities.

Historical Context

The Coleman Report was mandated by the Civil Rights Act of 1964. The act gave the U.S. Office of Education two years to produce a report that was expected to describe the inequality of educational opportunities in elementary and secondary education across the United States. Congress sought to highlight, particularly in the South, the differences between schools attended by whites and those attended by blacks (referred to as "Negroes," as was standard at the time).



The Coleman Report failed to accomplish one of the key goals that led Congress to commission the report: a forward march toward equal educational opportunity across racial groups. At the pace of the past 50 years, it will take roughly two and a half centuries to close the math achievement gap first seen in 1965.

But Congress, and the nation, got something very different from what most people expected. Working quickly as soon as the Civil Rights Act was signed into law, the Coleman research team drew a sample of over 4,000 schools, which yielded data on slightly more than 3,000 schools and some 600,000 students in grades 1, 3, 6, 9, and 12. The team asked students, teachers, principals, and superintendents at these schools a wide range of questions. The study broadened the measures of school quality beyond what policymakers envisioned. The surveys gathered objective information about "inputs," but they also asked about teacher and administrative attitudes and other subjective indicators of quality. The most novel aspect of the study was the assessment of students, who were given a battery of tests of both ability and achievement.

Coleman's team collected these data from schools across the country, tabulated them, analyzed them, and produced the mammoth report (and a second 548-page volume with descriptive statistics) within the two-year period. This dizzying pace of research is almost inconceivable at a time when high-speed computers were yet to become available.

The focus on hard, quantifiable facts cannot be overemphasized. It is difficult to find two consecutive pages in the report that do not contain at least one table or figure. In fact, it is easy to find 10 consecutive pages of dense tables or figures. As a result, a large portion of the potential readership was immediately bewildered by statistics, many of which were not commonly employed or broadly understood even within the academic community. It is exceedingly unlikely that more than a very few people actually read the entire report rather than relying on summaries or a sampling of the document's contents.

The difficulty of understanding the analysis and its implications was such that Daniel Patrick Moynihan organized a faculty seminar at Harvard that attracted some 80 researchers and met weekly for a year. Even among this erudite group, no clear consensus on what to make of the Coleman Report emerged. My own participation in this seminar as a graduate student set my entire career to the study of education policy.

A Summary

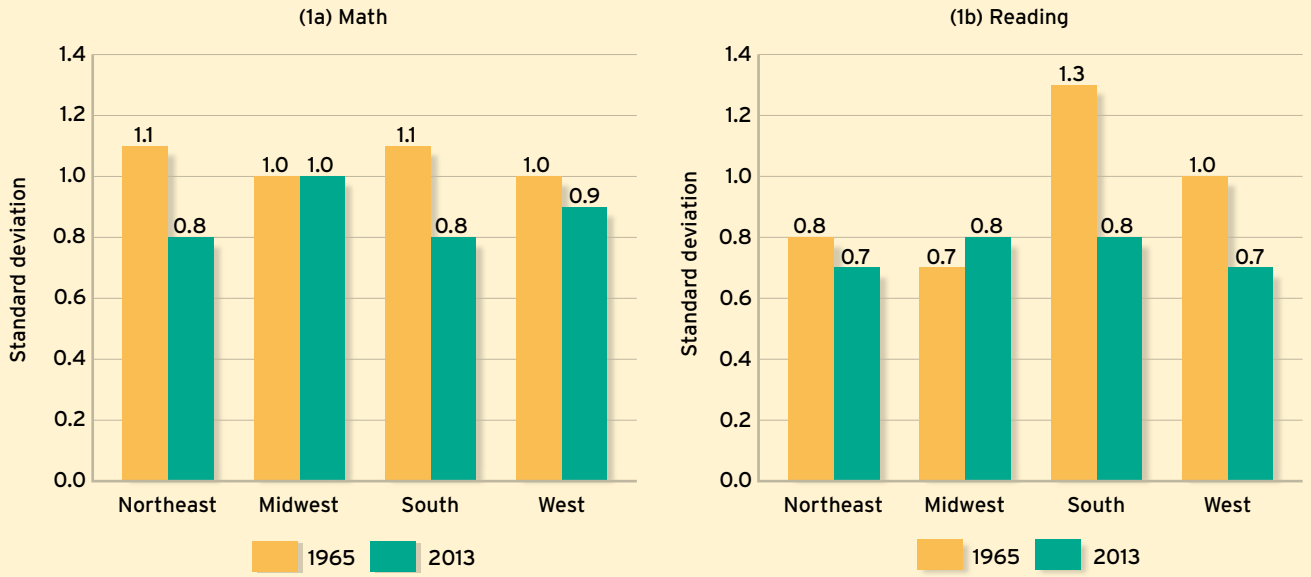
After 325 pages of charts, tables, and text, one gets to the enduring summary of the Coleman Report.

> Taking all these results together, one implication stands out above all: That schools bring little influence to bear on a child's achievement that is independent of his background and general social context; and that this very lack of an independent effect means that the inequalities imposed on children by their home, neighborhood, and peer environment are carried along to become the inequalities with which they confront adult life at the end of school.

Wrapped up in this statement are the ambiguities of meaning, the unclear translation into policies, and the inherent questions about analytical underpinnings that have persisted. For some, they point to the need for desegregation; for others, they suggest that schools do not matter; and for a third group, they highlight the overwhelming importance of the family.

The Black-White Achievement Gap Persists (Figure 1)

The black-white achievement gaps for 12th graders, based on the 2013 National Assessment of Educational Progress, remain unacceptably large, not much smaller than those identified by the Coleman Report for 1965. Progress in closing the gaps outside of the South has been largely disappointing.



NOTES: Data for 1965 combine the South and Southwest regions as "South." For the 2013 NAEP results, the following states are included in each region. Northeast: Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. Midwest: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin. South: Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia. West: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming. For 1965, Coleman included Arizona and New Mexico as part of the West, and Delaware, Maryland, and Washington, D.C., as part of the Northeast.

SOURCES: "Equality of Educational Opportunity" (1966), Table 3.121.3; National Assessment of Educational Progress, National Center for Education Statistics

One of Coleman's principal findings—often overlooked in the focus on the role of families, schools, and desegregation—was the shocking achievement disparities across races and regions within the United States. In 1965, Coleman tells us, the average black 12th grader in the rural South registered an achievement level that was comparable to that of a white 7th grader in the urban Northeast. That gap and other similar performance gaps never received the attention they deserved.

As a result, the Coleman report failed to accomplish one of the key goals that led Congress to commission the report in the first place: a forward march toward equal educational opportunity across racial groups. That simply happened haltingly in most parts of the country.

In both math and reading, the national test-score gap in 1965 was 1.1 standard deviations, implying that the average black 12th grader placed at the 13th percentile of the score distribution for white students. In other words, 87 percent of white 12th graders scored ahead of the average black 12th grader. What does it look like 50 years later? In math, the size of the gap has fallen nationally by 0.2 standard deviations, but that still leaves the average black 12th-grade student at only the 19th percentile of the white distribution. In reading, the achievement gap has improved slightly more than in math (0.3 standard deviations), but after a half century, the average black student scores at just the 22nd percentile of the white distribution.

As Figure 1 shows, the largest gains in both math and reading were found in the South, where the larger gaps observed in 1965 were brought in line with the rest of the nation by 2013. But the generally slow improvements in much of the rest of the country, including an expanded reading gap in the Midwest, attenuated the overall improvement.

After nearly a half century of supposed progress in race relations within the United States, the modest improvements in achievement gaps since 1965 can only be called a national embarrassment. Put differently, if we continue to close gaps at the same rate in the future, it will be roughly two and a half

centuries before the black-white math gap closes and over one and a half centuries until the reading gap closes. If "Equality of Educational Opportunity" was expected to mobilize the resources of the nation's schools in pursuit of racial equity, it undoubtedly failed to achieve its objective. Nor did it increase the overall level of performance of high school students on the eve of their graduation, despite the vast increase in resources that would be committed to education over the ensuing five decades (see Figures 2, 3, and 4).

Coleman did report a good deal of disparity in school resources from one part of the United States to the other, with the South lagging far behind the Northeast. But, within regions, racial differences in available resources were modest. Although it is difficult to make precise comparisons between then and now, the regional and racial disparities of today in education inputs are probably quite similar to those Coleman reported in 1966.

These overall descriptive findings can be taken as given, without quibbling over statistical methodology. But digging into the weeds of the Coleman Report, it is evident that the analysis of what determines achievement leaves much to be desired. The analysis has two major flaws. First, it attempts to assess what factors drive the observed differences in student achievement, but it does a poor job. Second, the approach fails to provide clear policy guidance on how achievement could be improved.



The finding in the Coleman Report that family-background factors powerfully affect student achievement is not and never has been disputed. Virtually all subsequent analyses have found measures of family background (parents' education, family structure, and so forth) to be a significant explanation of achievement differences.

In simplest terms, the statistical procedure of the Coleman Report relies on a problematic stepwise analysis of variance approach, which makes strong assumptions about which factors are fundamental causes of achievement and which are of secondary significance. Coleman assumed that family influences come first, and that school factors are to be introduced into the analysis only after all effects that can be attributed to the family are identified. Accordingly, the first step of the statistical analysis assesses how much of the achievement variation across schools could be attributed to variations in family background factors. Only after these background factors are fully accounted for is the second step taken—a look at the characteristics of the schools that make the biggest difference in determining the variation in student achievement.

This approach privileges family background over any indicators of school resources or peer group relationships, as it implicitly attributes all shared variation to those variables included in the first step of the stepwise modeling. For example, if parental education and teacher experience are both strongly related to achievement, and children from better-educated families attend schools with more-experienced teachers, then it will appear as if teacher experience has little effect while the effect of parental education is magnified. The first step, looking at just the relationship between achievement and parental education, actually incorporates both the direct effect of parental education on achievement and the indirect effect of the more-experienced teachers in their schools. When the analysis gets to the point of adding teacher experience to the explanation of achievement, the only marginal impact will come from the portion of variation in experience that is totally unrelated to family background.

But, more importantly, this partitioning of the variation in student achievement according to variations in underlying factors gives little indication of what could be expected from policies that alter the school inputs available to students. The statistical analysis relied exclusively on some crudely measured differences across schools, such as the number of days in the school year or the presence of a science lab. Most of their measures were not factors that would drive policy initiatives. Yet, the larger problem is that simply looking at the influence of the existing variation in these measures does not indicate the leverage on achievement that any would have. For example, the days in the school year showed relatively little variation, and, as such, variation in the length of school years could not explain much of the existing achievement variation, even if adding days to the school year would have a strong impact on achievement. Unfortunately, misinterpretations of these aspects of the Coleman analysis continue in the present day.

Among researchers with an understanding of the best ways

to estimate causal effects on educational achievement, none would rely on the methodology used by the Coleman team to estimate the effect of schools or teachers. The stepwise regression was problematic even in the 1960s and has been totally discredited as a method for estimating causal effects in the 50 years since.

Given this, I take the Coleman Report conclusions stated above as hypotheses, not as findings. What does current evidence say about these hypotheses?

Only Families Matter

That families have a strong, if not overwhelming, effect on student achievement is one of the most frequently repeated bumper-sticker claims of those who cite the Coleman Report. Analysts who claim that poverty explains the problems of the American school readily refer to Coleman as proof. The Economic Policy Institute's Richard Rothstein has declared that "the influence of social class characteristics is probably so powerful that schools cannot overcome it, no matter how well trained are their teachers and no matter how well designed are their instructional programs and climates." The campaign for a Broader, Bolder Approach to Education hints at the standard interpretation of the Coleman findings when it asserts that "poverty, which has long been the biggest obstacle to educational achievement, is more important than ever."

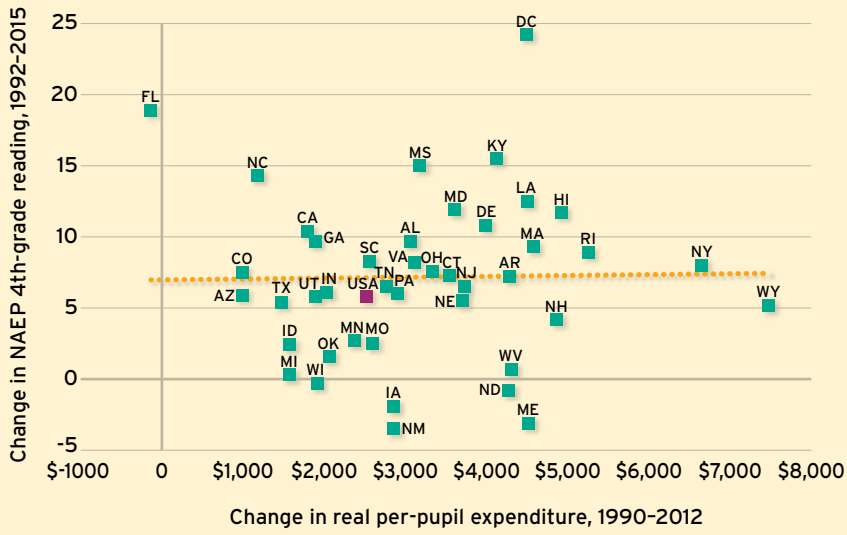
The Coleman Report itself measured family background by a series of survey questions given to the students that were combined into measures of urbanism, parents' education, structural integrity of the home, size of family, items in the home, reading material in the home, parents' interests, and parents' educational desires. Coleman did not measure family income, because he did not think students were a reliable source for this kind of information. Indeed, the word poverty appears just once in the entire report, in the summary; it was never used in the analysis. It is thus quite ironic that 21st-century references to Coleman regularly claim that he showed the major impact poverty had on student achievement.

Still, the finding that family-background factors powerfully affect student achievement is not and never has been disputed. Virtually all subsequent analyses have included measures of family background (education, family structure, and so forth) and have found them to be a significant explanation of achievement differences. Indeed, no analysis of school performance that neglects differences in family background can be taken seriously.

At the same time, the importance of this reality for education policy is quite unclear. Some argue that since poverty is strongly related to achievement, we must alleviate poverty before we can hope to have an effect of schools on achievement.

Achievement Gains Are Unrelated to Increased Spending (Figure 2)

There is no correlation between changes in real per-pupil spending over the past quarter century and changes in 4th-grade student achievement in reading, providing no reason to believe that increasing school spending will by itself boost student achievement.



SOURCE: National Assessment of Educational Progress, National Center for Education Statistics

For example, Diane Ravitch states that "[reformers believe] that schools can be fixed now and that student outcomes (test scores) will reach high levels without doing anything about poverty. But this makes no sense. Poverty matters." This type of interpretation of the Coleman Report and subsequent studies fails on several grounds.

Existing studies have generally accounted for family background by whatever measures were in their specific data set, ranging from family income to parental education to family structure to race and ethnicity. At some level, all of these measures are correlated with each other, and scholars are still not sure which is the "right measure." For example, some of the best research has focused on "family income" as a predictor of education success,

but Susan Mayer, a University of Chicago sociologist, has shown that unexpected changes in family income by themselves have little effect on a child's educational performance.

Moreover, the exact channels through which family resources have their impact on educational and lifetime successes remain uncertain. Is reading to the child decisive? Is the vocabulary of the parents? Is it the greater access to medical and dental services that children of more resourceful parents enjoy? Is it the more-sensitive child-rearing practices of the better-educated? Is it the greater interaction with adults that can occur in two-parent families that counts? Do more-resourceful parents find ways to place their children in more-effective educational settings? Most important, little evidence shows that just providing money to families can change the relevant family inputs, whatever they are.

Do Schools Matter?

Perhaps the largest long-term impact of the Coleman Report has been its effect on elite opinion about the contribution schools make to student achievement. The report's suggestion that schools add little beyond the family to student performance has provoked a bifurcated reaction. One side, which includes many school teachers and administrators, accepts this at face value, as it simply confirms what they already believe: schools should not be held responsible for poor student performance and achievement gaps that are driven by family background factors. The other side raises questions about the Coleman approach to estimating the relative importance of schools and families, and searches for other analytical methods and data sets that might open the question for further consideration.



How money is spent is much more important than how much is spent. Just providing more funds to a typical school district without any change in incentives and operating rules is unlikely to lead to systematic improvements in student outcomes.

The Coleman Report concludes that its measures of most school resources were only weakly associated with student achievement. Once family background and the nature of the peer group at school were taken into account, student achievement was unaffected by per-pupil expenditure, school size, the science lab facilities, the number of books in the library, the use of tracking by ability levels to assign students to classrooms, or other factors previously assumed to be indicators of what makes for a good school. In general, these findings have been reaffirmed by the scholarly community over the five decades since the report was written. Subsequent studies have found little in the way of systematic impacts of measured differences in resources among schools. On occasion, a specific study might find any one of these factors to be correlated with student performance, but, taken together, the vast proportion of results across a wide array of studies has found no statistically significant connection between the standard resources available to schools and the amount of learning taking place within the building.

Yet that is not the end of the story. While these findings appear clear, their interpretation calls for considerable care.

The Coleman data did not permit following the learning trajectories of individual students or looking at what happened within schools. Coleman tended to look at measures of school quality that administrators and policymakers rely on when defending their proposals to school boards. Those variables may not be correlated with student achievement, but that does not necessarily mean that schools are unimportant. It is quite possible that other, more-difficult-to-measure factors may be crucial for student learning.

Little attention was paid to indications in the Coleman Report that teachers might be a particularly critical school factor. But since the report's publication, scholars have developed more precise data on teacher effectiveness, and, by probing at differences in teacher quality within schools, have found very large impacts of teacher quality on student achievement. Admittedly, many teacher characteristics commonly used to measure teacher quality have little, if any impact on student performance. Whether teachers are certified, or obtain an advanced degree, or attend a specific college or university, or receive more or less mentoring or professional development turns out to be almost completely unrelated to a teacher's effectiveness in the classroom.

But measures of teacher effectiveness in the classroom (as estimated by the amount of learning taking place in classes under that teacher's supervision) do correlate with the learning taking place in that same teacher's classroom in subsequent years. In other words, qualitative differences among teachers have large impacts on the growth in student achievement, even though these differences are not related to the measured background characteristics or to the training teachers have received.

Scholars remain in the dark even today as to exactly why some teachers are effective (that is, why some teachers, year after year, have strong positive impacts on the learning of their pupils) while others are not. In short, it is easier to pick out good teachers once

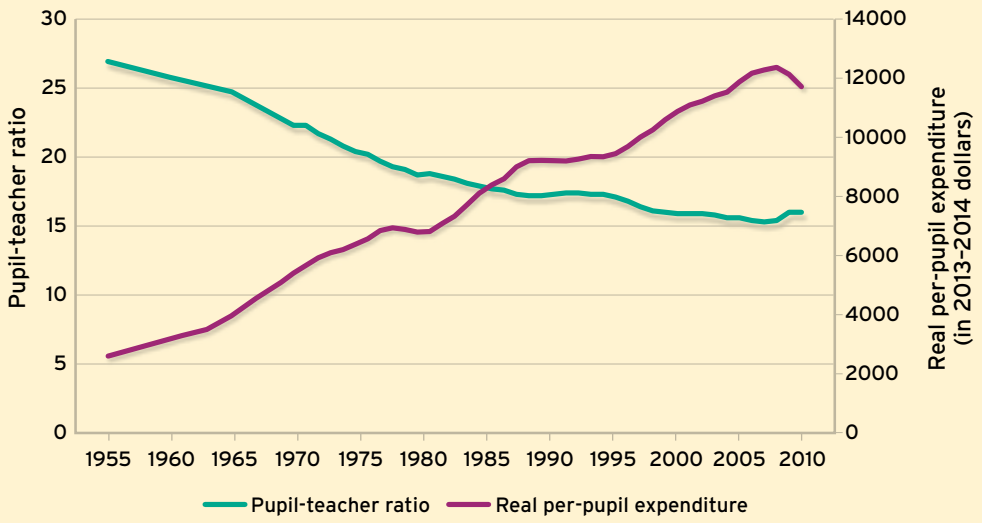
they have begun to teach than it is to train them or figure out exactly the secret sauce of classroom success.

Since most of the variation in teacher effectiveness is actually found within schools (i.e., between classrooms) and not between schools (Coleman's focus), the critical role of the teacher remained to be clearly documented by future scholars. For example, in work that I have done studying performance in disadvantaged urban schools, a top teacher can in one year produce an added gain from students of one full year's worth of learning compared to students suffering under a very ineffective teacher.

Stanford researcher Raj Chetty and his colleagues have shown that the effects of the teacher persist into adulthood. Those with the more-effective teacher will be more likely to pursue their education for a longer period of time and will earn more income by age 28.

More Teachers and More Money (Figure 3)

Since 1955, the number of students per teacher in public schools in the United States has fallen from 27 to 16, and per-pupil school spending has increased more than fourfold.



NOTE: Real per-pupil expenditures shown exclude capital outlay and debt service expenditures and are calculated based on average daily attendance.

SOURCE: National Center of Education Statistics

In short, research shows very large differences in teacher effectiveness. Moreover, variations in teacher effectiveness within schools appear to be much larger than variations between schools. Thus, the Coleman study failed to identify the importance of teacher quality and failed to grasp the policy relevance of within-school variation in teacher quality. These findings also illustrate vividly the problem introduced by the Coleman analytical approach: finding that measured teacher differences have limited ability to explain variations in student achievement is very different from concluding that schools and teachers cannot powerfully affect student outcomes.

Does Money Matter?

Coleman found that variations in per-pupil expenditure had little correlation with student outcomes. Although this was one of the key findings of the report, little attention was paid to this inconvenient fact. At the time, the Johnson administration was trumpeting a federally funded compensatory education program that was supposed to equalize educational opportunity by concentrating more funding on students living in low-income neighborhoods. But the finding gradually assumed greater importance in policy debates, as extensive subsequent research engendered by the Coleman Report reinforced this conclusion.

A defining moment came in the 1970s, when the California Supreme Court in *Serrano v. Priest* decided that in order to ensure equal educational opportunity for all children, all school districts in California must spend equal amounts per pupil, instigating a wave of school-finance court cases across the country. If expenditures must be equal in order for opportunity to be equal, then the amount spent per pupil must be critically important to student learning. Despite the Coleman findings, the claim that money matters was routinely made in courtrooms in nearly every state, provoking a bevy of research on the effects of school expenditure on student achievement. This is not the place to explore a debate that has relied on a mixture of scientific evidence, professional punditry, and misleading claims. Given the fiscal stakes involved, it is hardly surprising that the conversations have been politically charged and have led to an ongoing battle under the misleading sobriquet "money doesn't matter."

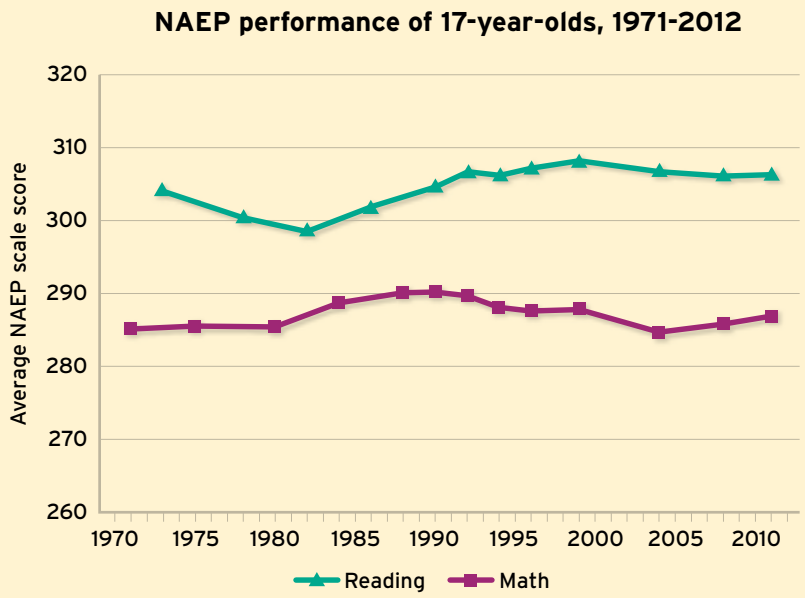
There remains the simple question as to whether, other things equal, just adding more money to schools will systematically lead to higher achievement. Figure 2 shows the overall record of states during the past quarter century. Changes in real state spending per pupil are uncorrelated with changes in 4th-grade student

achievement in reading. Similar results are obtained in math and in both math and reading at the 8th-grade level. Clearly, states have changed in many other ways than just expenditure, but there is no reason to conclude from these data that just providing money will by itself boost student achievement.

There now appears to be a general consensus that how money is spent is much more important than how much is spent. In other words, the research does not show that money never matters or that money cannot matter. But, just providing more funds to a typical school district without any change in incentives and operating rules is unlikely to lead to systematic improvements in student outcomes. That is what Coleman found, and that is what recent research says.

Stagnant Performance (Figure 4)

The performance in both reading and math of 17-year-olds on the National Assessment of Educational Progress today is virtually unchanged from that in the early 1970s.



SOURCE: National Assessment of Educational Progress, National Center for Education Statistics

Such a conclusion does not, however, resolve the question about the appropriate level of funding. Some argue that a certain level of funding is "necessary" even if not "sufficient" for improving student performance. Nonetheless, no research to date has defined the level that is necessary or adequate. Such efforts are continuously confounded by the fact that school funding is a rapidly moving target, as average U.S. spending on schools has quadrupled in real terms since 1960 (see Figure 3). Today, expenditures per pupil in the United States exceed those of nearly every other country in the world. Yet when it comes to student achievement, we see that U.S. student performance is virtually unchanged from that in the early 1970s (see Figure 4).

What remains to be unpacked is the precise ways in which expenditure needs to be directed and administered if it is to lift student achievement efficiently and effectively.

Lasting Impacts

The report's release quite dramatically changed the currency of policy debate to student outcomes. Prior to the report, school inputs—spending per pupil, teacher-pupil ratios, and the like—were customarily viewed as roughly synonymous with results. But both the approach and the conclusions of the Coleman Report altered this perspective.

The largest impact of the Coleman Report has been in the linkage of education research to education policy. It is difficult to find other areas of public policy where there is such a clear and immediate path from new research to the courts, to legislatures, and to policy deliberations. It is not unusual for research findings of working papers still with wet ink to be offered as proof that a new policy must be enacted.

There is, of course, a downside to this linkage. Often, policy research is cited when it gives the particular answer for which the policymaker is searching. As a result, there is a noticeable tendency on the part of many in the education policy world to cull the scientific literature for studies that come to a desired result. The Coleman Report has been twisted and turned in multiple ways by those who have a specific political agenda. Subsequent studies have suffered a similar fate.

Vastly more jarring is that the central goal of the report—the development of an education system that provides equal educational opportunity for all groups, and especially for racial minorities—has not been attained. Achievement gaps remain nearly as large as they were when Coleman and his team put pen to paper, even when better research has suggested ways to close them and even when policies have been promulgated that supposedly are explicitly designed to eliminate them.

Eric A. Hanushek is senior fellow at the Hoover Institution at Stanford University and research associate at the National Bureau of Economic Research.