



ILLUSTRATION / DAVE CUTLER

CONNECTING TO PRACTICE

HOW WE CAN PUT EDUCATION RESEARCH TO WORK

IN THE HALF CENTURY since James Coleman and his colleagues first documented racial gaps in student achievement, education researchers have done little to help close those gaps. Often, it seems we are content to recapitulate Coleman's findings. Every two years, the National Assessment of Educational Progress (a misnomer, as it turns out) reports the same disparities in achievement by race and ethnicity. We have debated endlessly and fruitlessly in our seminar rooms and academic journals about the effects of poverty, neighborhoods, and schools on these disparities. Meanwhile, the labor market metes out increasingly harsh punishments to each new cohort of students to emerge from our schools underprepared.

At the dawn of the War on Poverty, it was necessary for Coleman and his colleagues to document and describe the racial gaps in achievement they were intending to address. Five decades later, more description is unnecessary. The research community must find new ways to support state and local leaders as they seek solutions.

If the central purpose of education research is to identify solutions and provide options for policymakers and practitioners, one would have to characterize the past

by THOMAS J. KANE

five decades as a near-complete failure. There is little consensus among policymakers and practitioners on the effectiveness of virtually any type of educational intervention. We have learned little about the most basic questions, such as how best to train or develop teachers. Even mundane decisions such as textbook purchases are rarely informed by evidence, despite the fact that the National Science Foundation (NSF) and the Institute of Education Sciences (IES) have funded curricula development and efficacy studies for years.

The 50th anniversary of the Coleman Report presents an opportunity to reflect on our collective failure and to



Even mundane decisions such as textbook purchases are rarely informed by evidence, even though the federal government has funded curricula development and efficacy studies for years.

think again about how we organize and fund education studies in the United States. In other fields, research has paved the way for innovation and improvement. In pharmaceuticals and medicine, for instance, it has netted us better health outcomes and increased longevity. Education research has produced no such progress. Why not?

In education, the medical research model—using federal dollars to build a knowledge base within a community of experts—has manifestly failed. The What Works Clearinghouse (a federally funded site for reviewing and summarizing education studies) is essentially a warehouse with no distribution system. The field of education lacks any infrastructure—analogue to the Food and Drug Administration or professional organizations recommending standards of care—for translating that knowledge into local action. In the United States, most consequential decisions in education are made at the state and local level, where leaders have little or no connection to expert knowledge. The top priority of IES and NSF must be to build connections between scholarship and decisionmaking on the ground.

Better yet, the federal research effort should find

ways to embed evidence gathering into the daily work of school districts and state agencies. If the goal is to improve outcomes for children, we must support local leaders in developing the habit of piloting and evaluating their initiatives before rolling them out broadly. No third-party study, no matter how well executed, can be as convincing as a school's own data in persuading a leader to change course. Local leaders must see the effectiveness (or ineffectiveness) of their own initiatives, as reflected in the achievement of their own students.

Instead of funding the interests and priorities of the academic community, the federal government needs to shift its focus toward enabling researchers to support a culture of evidence discovery within school agencies.

An Evolving Understanding of Causality

In enacting the Civil Rights Act of 1964, Congress mandated a national study on racial disparities in educational opportunity—giving Coleman and his colleagues two years to produce a report. The tight deadline allowed no time to collect baseline achievement data and follow a cohort of children. Moreover, the team had neither the time nor the political mandate to assign groups of students to specific interventions in order to more thoroughly identify causal effects. Their only recourse was to use cross-sectional survey data to try to identify the mechanisms by which achievement gaps are produced and, presumably, might be reversed.

Given the time constraints, Coleman used the proportion of variance in student achievement associated with various educational inputs—such as schools, teacher characteristics, student-reported parental characteristics, and peer characteristics—as a type of divining rod for identifying promising targets for intervention. His research strategy, as applied to school effects, is summarized in the following passage from the report:

The question of first and most immediate importance to this survey in the study of school effects is how much variation exists between the achievement of students in one school and those of students in another. *For if there were no variation between schools in pupils' achievement, it would be fruitless to search for effects of different kinds of schools upon achievement* [emphasis added].

In other words, Coleman's strategy was to study how much the achievement of African American and white

students varied depending on the school they attended, and then use that as an indicator of the potential role of schools in closing the gap.

This strategy had at least three flaws: First, if those with stronger educational supports at home and in society were concentrated in certain schools, the approach was bound to overstate the import of some factors and understate it for others. It may not have been the schools, but the students and social conditions surrounding them that differed.

Second, even if the reported variance did reflect the causal effects of schools, the approach confuses prevalence with efficacy. Suppose there existed a very rare but extraordinarily successful school design. Schools would still be found to account for little of the variance in student performance, and we would overlook the evidence of schools as a lever for change. Given that African Americans (in northern cities and in the South) were just emerging from centuries of discrimination, it is unlikely that any 1960s school systems would have invested in school models capable of closing the achievement gap.

Third, the percentage-of-variance approach makes no allowance for “bang for the buck” or return on investment. Different interventions—such as new curricula or class-size reductions—have very different costs. As a result, within any of the sources of variance that Coleman studied, there may have been interventions that would have yielded social benefits of high value relative to their costs.

The fact that some of Coleman’s inferences have apparently been borne out does not mean that his analysis was ever a valid guide for action. (Even a coin flip will occasionally yield the right prediction.) For instance, because there was greater between-school variance in outcomes for African American students than for white students (especially in the South), Coleman concluded that black students would be more responsive to school differences. At first glance, Coleman’s original interpretation seems prescient: a number of studies—such as the Tennessee STAR experiment—have found impacts to be larger for African American students. However, such findings do not validate his method of inference. The between-school differences in outcomes Coleman saw might just as well have been due to other factors, such as varying degrees of discrimination in the rural South.

While there were instances where Coleman “got it right,” in other cases his percentage-of-variance metric pointed in the wrong direction. For example, in the 1966 report, the between-school variance in student test scores was larger for verbal skills and reading compre-

If the central purpose of education research is to identify solutions and provide options for policymakers and practitioners, one would have to characterize the past five decades as a near-complete failure.

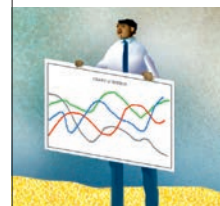
hension than for math. Coleman’s reasoning would have implied that verbal skills and reading would be more susceptible to school-based interventions than math would. However, over the past 50 years, studies have often shown just the opposite to be true. Interventions have had stronger effects on math achievement than on reading comprehension.

It was not until 2002, 36 years after the Coleman Report, that the education research enterprise finally began to adopt higher standards for inferring the causal effects of interventions. Beginning with the leadership of Russ Whitehurst at the Institute of Education Sciences, IES has begun shifting its grants and contracts away from correlational studies like Coleman’s and toward those that evaluate interventions with random-assignment and other quasi-experimental designs.

As long as that transition toward intervention studies continues, perhaps it is just a matter of time before effective interventions are found and disseminated.



Beginning with the leadership of Russ Whitehurst, IES has begun shifting its grants and contracts toward those that evaluate interventions with random-assignment and other quasi-experimental designs.



However, I am not so confident. The past 14 years have not produced a discernible impact on decisionmaking in states and school districts. Can those who argue for staying the course identify instances where a school district leader discontinued a program or policy because research had shown it to be ineffective, or adopted a new program or policy based on a report in the What Works Clearinghouse? If such examples exist, they are rare.

Federal Funds for Education Research

The Coleman Report is often described as “the largest and most important educational study ever conducted.” In fact, the 1966 study cost just \$1.5 million, the equivalent of \$11 million today. In 2015, the combined annual budget for the Institute of Education Sciences (\$578 million) and the education research conducted by the National Science Foundation (\$220 million) was equivalent to the cost of 70 Coleman Reports. Much more of that budget should be used to connect scholarship with practice and to support a culture of evidence gathering within school districts and state agencies.

As illustrated in Figure 1, the budget for the Institute of Education Sciences is allocated across four national centers. The largest is the National Center for Education Statistics (NCES), with a total annual budget of \$278 million. Roughly half of that amount (\$140 million) pays for the National Assessment of Educational Progress, which provides a snapshot of achievement nationally and by state and urban district. Most of the remainder of the NCES budget goes to longitudinal surveys (such as the Early Childhood Longitudinal Study of the kindergarten class of 2011 and the High School Longitudinal Study of 2009) and cross-sectional surveys (such as the Schools and Staffing Survey and National Household Educational Survey). Those surveys are designed and used by education researchers, primarily for correlational studies like Coleman’s.

The National Center for Education Research (NCER) is the second-largest center, with an annual budget of \$180 million. NCER solicits proposals from researchers at universities and other organizations. In 2015, NCER received 700 applications and made approximately 120 grants. Proposals are evaluated by independent scholars in a competitive, peer-review process. NCER also funds postdoctoral and predoctoral training programs for education researchers. Given its review process, NCER’s funding priorities tend to reflect the interests of the academic community. The National Center for Special Education Research (NCSER), analogous to NCER, funds studies on special education through solicited grant programs.

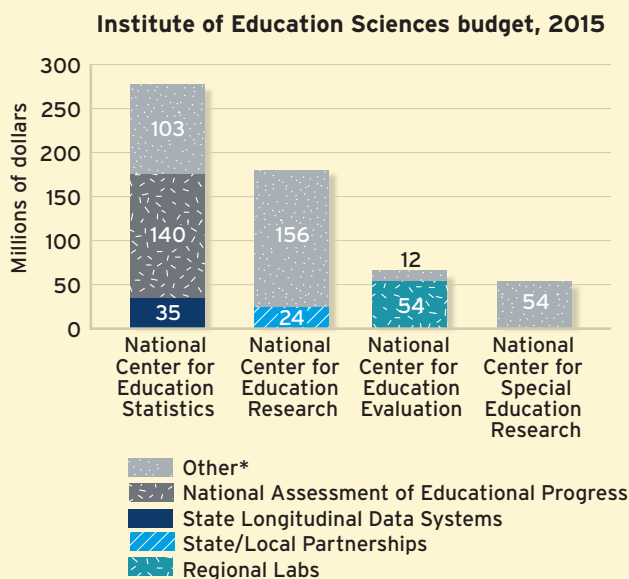
The National Center for Education Evaluation and Regional Assistance (NCEE) manages the Education Resources Information Center (ERIC)—an online library of research and information—and the What Works Clearinghouse. NCEE also funds evaluation studies of federal initiatives and specific interventions, such as professional development efforts. In principle, NCEE could fund evaluation studies for any intervention that states or districts might use federal funds to purchase.

The Disconnect between Research and Decisionmaking

While the federal government funds the lion’s share of education research, it is state and local governments that make most of the consequential decisions on such matters as curricula, teacher preparation,

Research Budget Breakdown (Figure 1)

The \$578 million budget of the Institute of Education Sciences is allocated across four national centers. Only \$113 million—less than 20 percent—is devoted to local, state, and regional programs.



*Includes research grants, evaluation studies, training programs, survey data collection, and online databases

SOURCE: U.S. Department of Education

teacher training, and accountability. Unfortunately, the disconnect between the source of funding and those who could make practical use of the findings means that the timelines of educational evaluations rarely align with the information needs of the decisionmakers (for instance, the typical evaluation funded by NCEE requires six years to complete). It also means that researchers, rather than policymakers and practitioners, are posing the questions, which are typically driven by debates within the academic disciplines rather than the considerations of educators. This is especially true at NCER, most of whose budget is devoted to funding proposals submitted and reviewed by researchers. At NCES as well, the survey data collection is guided by the interests of the academic community. (The NAEP, in contrast, is used by policymakers and researchers alike.)

As mentioned earlier, fields such as medical and pharmaceutical research have mechanisms in place for connecting evidence with on-the-ground decisionmaking. For instance, the Food and Drug Administration uses the evidence from clinical trials to regulate the availability of pharmaceuticals. And professional organizations draw from experts' assessment of the latest findings to set standards of care in the various medical specialties.

To be sure, this system is not perfect. Doctors regularly prescribe medications for "off-label" uses, and it often takes many years for the latest standards of care to be adopted throughout the medical profession. Still, the FDA and the

medical organizations do provide a means for federally funded studies to influence action on the ground.

Education lacks such mechanisms. There is no "FDA" for education, and there never will be. (In fact, the 2015 reauthorization of the Elementary and Secondary Education Act reduces the role of the federal government and returns power to states and districts.)

In 2015, the combined annual budget for the Institute of Education Sciences and the education research spending by the National Science Foundation was equivalent to the cost of 70 Coleman Reports. Much more of that budget should be used to connect research with practice.

Professional organizations of teachers, principals, and superintendents focus on collective bargaining and advocacy, not on setting evidence-based professional standards for educators.

By investing in a central body of evidence and building a network of experts across a range of research topics, such as math or reading instruction, IES and NSF are mimicking the medical model. However, the education research enterprise has no infrastructure for translating that expert opinion into local practice.

To be fair, IES under its latest director, John Easton, was aware of the disconnect between scholarship, policy, and practice and attempted to forge connections. NCES, NCER, and NCEE all provide some amount of

support for state and local efforts, as Figure 1 highlights. For instance, the majority of NCEE's budget (\$54 million out of a total of \$66 million) is used to fund 10 regional education labs around the country. Each of the labs has a governing board that includes representatives from state education agencies, directors of research and evaluation from local school districts, school superintendents, and school board members. However, the labs largely operate outside of the day-to-day workings of state and district agencies. New projects are proposed by the research firms holding the contracts and must be approved in a peer review process. For the most part, the labs are not building the




In contrast to the field of education, the Food and Drug Administration and medical organizations provide a means for federally funded research to influence action on the ground.



capacity of districts and state agencies to gather evidence and measure impacts but are launching research projects that are disconnected from decisionmaking.

While NCEE is at least trying to serve the research needs of state and local government, less than 13 percent of the NCES and the NCER budgets is allocated to state and local support. NCES does oversee the State Longitudinal Data Systems (SLDS) grant program. With funding from that program, state agencies have been assembling data on students, teachers, and schools, and linking them over time, making it



Now that we know that some school-based interventions can shrink the achievement gap, we need researchers to collaborate with school districts around the country to develop, test, and scale up the promising ones.

possible to measure growth in achievement. Those data systems—developed only in the past decade—will be vital to any future effort by states and districts to evaluate programs and initiatives. However, the annual budget for the SLDS program is just \$35 million of the total NCES budget of \$278 million. For the moment, the state longitudinal data systems are underused, serving primarily to populate school report-card data for accountability compliance.

NCER sets aside roughly \$24 million per year for the program titled Evaluation of State and Local Programs and Policies, under which researchers can propose to partner with a state or local agency to evaluate an agency initiative. Such efforts are the kind that IES should be supporting more broadly. However, because the program is small and it is scholars who know the NCER application processes, such projects tend to be initiated by them rather than by the agencies themselves. It is not clear how much buy-in they have from agency leadership.

A New Emphasis on State and Local Partnerships

IES must redirect its efforts away from funding the interests and priorities of the research community and toward building an evidence-based culture within districts and state agencies. To do so, IES needs to create tighter connections between academics and decisionmakers at the state and local levels. The objective should be to make it faster and cheaper (and, therefore, much more common) for state and local leaders to

pilot and evaluate their initiatives before rolling them out broadly.

Taking a cue from NCER's program for partnerships with state and local policymakers, IES should offer grants for researchers to evaluate pilot programs in collaboration with such partners. But to ensure the buy-in of leadership, state and local governments should be asked to shoulder a small portion (say, 15 percent) of the costs. In addition, one of the criteria for evaluating proposals should be the demonstrated commitment of other districts and state agencies to partici-

participate in steering committee meetings. Such representation would serve two purposes: it would increase the likelihood that promising programs could be generalized to other districts and states, and it would lower the likelihood that negative results would be buried by the sponsoring agency. As the quality and number of such proposals increased, NCER could real-

locate its research funding toward more partnerships of this kind.

However, if the goal is to reach 50 states and thousands of school districts, our current model of evaluation is too costly and too slow. If it requires six years and \$12 million to evaluate an intervention, IES will run out of money long before the field runs out of solutions to test. We need a different model, one that relies less on one-time, customized analyses. For instance, universities and research contractors should be asked to submit proposals for helping state agencies and school districts not just in evaluating a specific program but in building the capacity of school agencies for piloting and evaluating initiatives on an ongoing basis. The state longitudinal databases give the education sector a resource that has no counterpart in the medical and pharmaceutical industries. Beginning with the No Child Left Behind Act of 2001, U.S. students in grades 3 through 8 have been tested once per year in math and English. That requirement will continue under the 2015 reauthorization bill, the Every Student Succeeds Act. Once a set of teachers or students are chosen for an intervention, the state databases could be used to match them with a group of students and teachers who have similar prior achievement and demographic characteristics and do not receive the intervention. By monitoring the subsequent achievement of the two groups, states and districts could gauge program impacts more quickly and at lower cost. The most promising interventions could later be confirmed with randomized field trials. However, recent studies using

randomized admission lotteries at charter schools and the random assignment of teachers has suggested that simple, low-cost methods, when they control for students' prior achievement and characteristics, can yield estimates of teacher and school effects that are similar to what one observes with a randomized field trial. Perhaps nonexperimental methods will yield unbiased estimates on other interventions as well.

In addition, IES should invite universities and research firms to submit proposals for convening state legislators, school board members, and other local stakeholders to learn about existing data on effective and ineffective programs in a particular area, such as preschool education or teacher preparation. IES should experiment with a range of strategies to engage with state and local agencies, and as effective ones are found, more of its budget should be allocated to such efforts.

Although the federal government provides the lion's share of research funding in education, state and local governments make a crucial contribution. Until recently, the primary costs of many education studies—including the Coleman Report—derived from measuring student outcomes and, in the case of longitudinal studies, hiring survey research firms to follow students and teachers over time. With the states investing \$1.7 billion annually on their assessment programs, much of that cost is now borne by states and districts.

Reason for Optimism

Fifty years after the Coleman Report, racial gaps in achievement remain shamefully large. Part of the blame rests with the research community for its failure to connect with state and local decisionmakers. Especially now that the federal government is returning power to states under Every Student Succeeds, federal efforts should be refocused to more effectively help states and districts develop and test their initiatives. The stockpiles of data on student achievement accumulating within state agencies and districts offer a new opportunity to engage with decisionmakers. Local leaders are more likely to act based on findings from their own data than on any third-party report they may find in the What Works Clearinghouse. If the research community were to combine IES's post-2002 emphasis on evaluating interventions with more creative strategies for engaging state and local decisionmakers, U.S. education

could begin to make more significant progress.

There is reason for optimism. Indeed, the Coleman Report's conclusion that schools had little hope of closing the achievement gap has been proven unfounded. In recent years, several studies using randomized admission lotteries have found large and persistent impacts on student achievement, even for middle school and high school students. For instance, students admitted by lottery to a group of charter schools in Boston increased their math achievement on the state's standardized test by 0.25 standard deviations per year



Especially now that the federal government is returning power to states under the Every Student Succeeds Act, signed into law by President Obama on December 10, 2015, federal research efforts should be refocused to more effectively help states and districts develop and test their initiatives.

in middle school and high school. Large impacts were also observed on the state's English test: 0.14 standard deviations per year in middle school and 0.27 standard deviations per year in high school. Similarly, a Chicago study of an intensive math-tutoring intervention with low-income minority students in 9th and 10th grades suggested impacts of 0.19 to 0.31 standard deviations—closing a quarter to a third of the achievement gap in one year. Now that we know that some school-based interventions can shrink the achievement gap, we need scholars to collaborate with school districts around the country to develop, test, and scale up the promising ones. Only then will we succeed in closing the gaps that Coleman documented 50 years ago.

Thomas J. Kane is the Walter H. Gale Professor of Education and faculty director of the Center for Education Policy Research at Harvard University.