

1

Bridging the Achievement Gap

JOHN E. CHUBB AND TOM LOVELESS

Overstating the importance of the achievement gap is not easy. The difference in educational achievement between white students, on the one hand, and African American and Hispanic students, on the other, is large and persistent. In the last decade it has gotten worse. The average black or Hispanic student, in elementary, middle, or high school, currently achieves at about the same level as the average white student in the lowest quartile of white achievement. These differences have dire consequences once students leave school. Blacks and Hispanics are much less likely than whites to graduate from high school, acquire a college or advanced degree, or earn a living that places them in the middle class. Blacks and Hispanics are much more likely than whites to suffer the social problems that often accompany low income. If the achievement gap could be reduced, the fortunes of blacks and Hispanics would not only be raised, but the social and economic differences that intensify the country's racial tensions would also be ameliorated.

The achievement gap has been extensively documented. In 1966, in the first truly comprehensive examination of student achievement in the United States, a team of federally sponsored researchers, led by the eminent sociologist James Coleman, found a large black-white achievement gap—which it attributed primarily to family backgrounds, of the students themselves and of their classmates. In 1970 the federal government launched the National Assessment of Educational Progress (NAEP), a random testing of American

students every two to four years and a means, for the first time, to track student achievement nationwide. These tests showed black and Hispanic students to be roughly four years behind white students on average by age seventeen. In addition, the 1970s saw large numbers of blacks and Hispanics seeking college admission and scoring approximately a standard deviation behind whites on the SAT.

Over the ensuing years, data from the NAEP, the SAT, and increasingly common state and district testing programs continued to show a racial gap in achievement. Scholars debated its causes—family, peer groups, culture, discrimination, heredity, and schooling. Policymakers lamented its intractability. Gains made by blacks and Hispanics from the late 1970s to the late 1980s on national measures eroded in the 1990s. And these declines occurred during perhaps the most aggressive era of school reform in the nation's history. A fair conclusion, as the definitive volume on the subject (*The Black-White Test Score Gap*, edited by Christopher Jencks and Meredith Phillips) reached in 1998, is that far more is known about the nature of the achievement gap—its causes and its consequences—than about how to fix it.

The situation, however, may finally be changing. Achievement trends, discouraging as they may be, do not give a complete picture of developments in student achievement. Around the country a number of specific efforts are showing that the achievement gap can be bridged. Schools and school reforms are boosting the achievement of Hispanic and, especially, African American students to levels nearing those of whites. Disparate in approach and involving relatively few students, these efforts are nonetheless important. They are potentially replicable. They offer lessons that might be learned and applied widely. They offer hope that the achievement gap might one day, soon, be reduced meaningfully. This book provides, for the first time in one place, the evidence of these efforts—evidence that when taken together is remarkably encouraging.

Over a decade ago, the state of Tennessee launched one of the most important experiments in education reform to investigate the effects of class size on student achievement. Project STAR was designed as a large-scale experiment with random assignment of students to treatment and control groups; that is, to classes of regular or small size or regular size with a teacher's aide. It represented the most powerful of research designs and a rarity in public policy—everyone eligible for a new policy benefit (in this case, a smaller class) wants or expects to receive it. Class-size reduction also happens to be one of the more popular, albeit expensive, reforms being implemented in school systems throughout America today.

In their chapter, “Would Smaller Classes Help Close the Black-White Achievement Gap?” Alan B. Krueger and Diane M. Whitmore put the data from Project STAR to what might be considered the ultimate test. The students in the project, who were in grades K–3 when they experienced different sized classes, have now either finished high school or dropped out. Krueger and Whitmore explore the long-term effects of smaller classes on student achievement. They found that, while all types of students enrolled in smaller classes achieved at higher levels than students enrolled in regular classes (with or without aides) during the years they were in grades K–3, the achievement advantage of smaller classes largely disappeared for white students once they returned to regular-size classes in fourth grade. For black students, however, the achievement advantage diminished in grade four but then stabilized at about 5 national percentiles over the long haul, reducing the black-white achievement gap for those students by 15 percent. What is more, the advantages of smaller classes showed up in participation rates of taking the SAT and the American College Test (ACT). Black students enrolled in smaller classes in their primary years took college entrance exams, years later, at an 8 percent higher rate than black students in regular classes. This amounts to more than a 50 percent reduction in the black-white gap in college entrance exam participation—and a potentially significant reduction in differences in college attendance rates for blacks and whites.

These differences are to be taken seriously because they are enduring products of well-designed experimental interventions. But class size is not the only reform for which experimental evidence is becoming available. School vouchers, a reform of arguably the greatest potential significance, have been implemented through randomized field trials in multiple settings over the last five years. Vouchers, which entitle students to attend any public or private school willing to accept the voucher for all or part of the cost of attendance, could stimulate dramatic change in America’s schools. Poorly performing schools could easily lose all of their students and revenue as students choose to use their vouchers to go elsewhere. Schools offering quality education, whether public or private, existing or entirely new, would flourish. Vouchers could bring about change not only in the practices of schools, as they seek to maintain or build enrollment, but also in the mix of schools in existence.

The effects of vouchers are hardly a simple matter to predict. Families could make poor choices. Poor families and poorly educated families, including the families of many black and Hispanic youngsters, could make the worst choices. Schools could cater to families without improving education.

The marketplace might provide better schooling for some but worse schooling for others, and, as a result, the rich could get richer and the poor could get poorer. These issues have been the most vigorously debated in education reform—in part, because they deal with the fundamental organizing principles of public education, but in part because so little direct evidence exists to resolve them.

Paul E. Peterson and William G. Howell, in “Voucher Programs and the Effect of Ethnicity on Test Scores,” review the most direct evidence of how vouchers might work. Three state governments—Florida, Ohio, and Wisconsin—and private philanthropists in several major cities currently provide vouchers for students from low-income families. In several of these programs, students are awarded vouchers through a lottery, creating the opportunity to study the effects of the voucher experimentally. Students using vouchers can be compared with students who sought but failed to win vouchers. Peterson and Howell carry out these comparisons over two years in Dayton, Ohio, New York City, and Washington, D.C., where the vouchers are provided by philanthropy. Their key finding is that black students show a clear benefit from attending private school as compared with black students attending public school. The same is not true for white or Hispanic students. Their achievement is no higher in private school than in public. While the effect is not identical for blacks across all cities, it averages 6.3 percentiles, a full third of the black-white test score gap in these data and a clear sign that underachievement by blacks may have much to do with the schools black students attend.

Whether a school is public or private, understanding what schools precisely do to benefit students is important. Over the last decade, reformers have attempted to specify more thoroughly the elements of effective schooling, and they have produced in the process a number of comprehensive reform models. Among these are the Modern Red School House, Direct Instruction, America’s Choice, Edison Schools, and, the most widely adopted, Success for All. These models go well beyond changing curriculum or instruction, the traditional domains of school reform. They attempt to change school practice from how they organize to how they teach to how they work with families to, in some cases, everything schools do. All of these models share one other thing in common: they have been adopted primarily by schools serving economically disadvantaged students and students of color.

Robert E. Slavin and Nancy A. Madden are the architects of Success for All, which was created in 1987 and is used today by nearly one million students in forty-eight states and eighteen hundred schools, virtually all with

high levels of poverty. In “‘Success for All’ and African American and Latino Student Achievement,” Slavin and Madden review the effects of Success for All on reading achievement in several different settings. First, they consider thirty-five schools, including some six thousand students, where matched control schools could provide reliable measures of program effects. In these thirty-five schools, which are majority African American or Hispanic, Success for All schools achieved about 0.5 standard deviations higher at each elementary grade level than control schools and were one full grade level ahead of the controls by fifth grade. Second, Slavin and Madden review an independent study of Success for All and other reform programs in Memphis, Tennessee, where students are nearly all African American. (The independent study was conducted by famed school accountability expert and scholar William Sanders.) The students in Success for All made the top gains in the city. Third, a study of every school in the state of Texas using Success for All showed black students exceeding the statewide improvement rate on the Texas Assessment of Academic Skills (TAAS), the state’s accountability test, by 50 percent. Hispanic gains were better, too, but by lesser amounts.

If Success for All is showing in large numbers how the achievement of traditionally low-achieving groups can be moved forward, other programs are showing in small numbers but just as dramatically what a difference excellent schooling can make. Alex Molnar and a team of researchers from the University of Wisconsin–Milwaukee investigate Wisconsin’s Student Achievement Guarantee in Education (SAGE) program. Beginning in 1995, Wisconsin offered schools serving large numbers of poor children an extra \$2,000 for each low-income student. The money was to be spent on four reforms: (1) reducing pupil-teacher ratios in kindergarten through third grade, (2) keeping campuses open from early morning to late evening, (3) implementing a rigorous curriculum, and (4) providing professional development activities to instructional staff. Soon after the program began, class-size reduction became the focal point of SAGE in most of the participating schools, and the program expanded from 30 schools the first year to 566 schools in 2001.

The researchers, in “Wisconsin’s SAGE Program and Achievement through Small Classes,” analyze the test scores of SAGE students as they moved from first through third grades, comparing their achievement with non-SAGE students at similar schools. The findings are positive for SAGE. In reading, math, and language arts, SAGE students gained more than the comparison group in first and third grades—and registered approximately equal gains in second grade. Moreover, after three years in the programs, even though significant gains were evidenced by all racial groups, African American students gained more than whites. The black-white gap narrowed by

about a third without white students losing ground. In the non-SAGE comparison schools, the achievement gap slightly expanded. In terms of finding politically feasible solutions to bridging the achievement gap, this is an important development. Policies that are perceived to close gaps at the expense of majority students—busing in the 1970s may be the clearest example—are destined to draw strong political opposition. Thus, Wisconsin may have developed a program that is both educationally sound and politically attractive.

Texas is another state that has targeted the achievement gap. Laurence A. Toenjes and a team of researchers from the University of Houston analyze the impact of the Texas testing and accountability system, in “High-Stakes Testing, Accountability, and Student Achievement in Texas and Houston.” In the mid-1980s, Texas was a pioneer in designing a high-stakes accountability system. Several incentives were introduced, from students being required to pass tests before they could graduate from high school or participate in team sports to individual schools being rewarded or sanctioned based on students’ test scores. The Texas Assessment of Academic Skills is a battery of criterion-referenced tests serving as the linchpin of the system. To promote equity, Texas was one of the first states to require that schools annually report disaggregated TAAS scores, allowing for the monitoring of progress by each major ethnic and racial group. That requirement was incorporated into the Leave No Child Behind Act, the federal education legislation signed by President George W. Bush in 2002.

The Houston researchers document how all three of Texas’s major ethnic groups have registered solid academic gains on the TAAS. For the state as a whole, 55.6 percent of students passed the tests in 1994. In 1999 the passing rate had grown to 78.3 percent. The improvement was most pronounced for Hispanics and African Americans. The Hispanic passing rate jumped from 41.1 percent to 70.1 percent. The passing rate of African Americans nearly doubled, from 33.3 percent to 64.0 percent. The passing rate for whites improved from 69.4 percent to 87.9 percent. In recent years, critics have questioned whether the TAAS gains are real. The researchers address these concerns by analyzing TAAS data from several different angles—using mean scores instead of passing rates, comparing cohorts, and focusing only on Houston. They also evaluate enrollment data to determine whether an increase in student dropout rates may have inflated the TAAS scores. They conclude that the gains are real and that the achievement gap has narrowed substantially in Texas.

In their chapter, “Schools That Work,” Abigail Thernstrom and Stephan Thernstrom offer a detailed look into three schools, one in Los Angeles, one

in Houston, and one in New York City's borough of the Bronx, that are driving student achievement to eye-popping gains. The three schools are the progenitor (Los Angeles) and first working models of the Knowledge Is Power Program, or KIPP. These schools have taken students from black, Hispanic, and highly disadvantaged backgrounds and raised them from the lowest quartile of the national achievement distribution to the top quartile. These schools have made attendance at competitive universities a reality for students who would normally not attend college at all.

The Thernstroms argue that despite the bleak history of the achievement gap, which they review in detail, there is reason for optimism. No evidence exists that past school reform has attempted to create schools with the vision, determination, and fundamentally sound practice found at KIPP. Yet, if schools could put into place what the founders of KIPP first experienced with a special teacher in Los Angeles and then established in public schools in Houston and the Bronx in the mid-1990s, then clearly, as the tired conviction of educators goes, "all children can learn." The challenge is taking the successes that can be so great in a few schools and carrying them over to many schools.

David Klein, in "High Achievement in Mathematics: Lessons from Three Los Angeles Elementary Schools," identifies a weak curriculum as a major source of inequality in achievement and a strong curriculum as a means of closing gaps. Klein describes three elementary schools in the Los Angeles area: Bennett-Kew, William H. Kelso, and Robert Hill Lane. The schools serve predominantly disadvantaged students. At Bennett-Kew, for example, 77 percent of students qualify for free or reduced-price lunch, 51 percent are African American, 48 percent are Latino, and 29 percent are limited in English proficiency. This demographic profile is, regrettably, one that many people automatically associate with low test scores. Yet on California's annual academic assessments, the three schools consistently score at levels typical of schools in the state's most privileged neighborhoods.

The three elementary schools mainly resist faddish, cutting-edge approaches to teaching and relentlessly focus on teaching the core skills of reading and mathematics. Teachers instruct students on this knowledge and test regularly to make certain that pupils have learned it. In other words, the teachers in these schools teach students solid content. They do not facilitate, guide, or explore. Whole language is out; phonics is in. Calculators and National Council of Teachers of Mathematics math reform are out; arithmetic is in. Process approaches to writing are out; grammar, spelling, and punctuation are in. Bilingual education is out; English immersion and class assistants who help students develop facility with the English language are in.

The schools' principals—Nancy Ichinaga, Marjorie Thompson, and Sue Wong—fought hard to establish and maintain an academic focus at their schools and, by cultivating close ties with parents, have built an ethos of achievement from the classroom to the home. Klein concludes that clear, high-quality standards, textbooks and learning materials that convey high expectations to students, and teachers with sufficient content knowledge to make students reach beyond what they currently know are the key ingredients to school success—and to boosting the achievement of poor and minority students.

Samuel R. Lucas and Adam Gamoran examine the controversial practice of tracking in high school, in "Tracking and the Achievement Gap." They compare data on high school sophomores from two national surveys—High School and Beyond, conducted in 1980, and the National Education Longitudinal Study, conducted in 1990. Critics of tracking have long charged that the practice treats minority students unfairly, channeling them into low-track classes in which very little learning takes place. Lucas and Gamoran are interested in how changes in students' course taking in the 1980s affected achievement gaps and in assessing whether different track indicators—the manner in which a student's track location is reported—alter the measurement and interpretation of tracking's effects.

Did changes in tracking in the 1980s affect achievement gaps? The benefit of taking college-track courses grew during the decade. With all racial groups enrolling in more rigorous, college prep courses in 1990—African Americans, in particular—the black-white gap was cut in half. In 1980 Hispanic students were at a disadvantage in enrolling in college prep classes. However, by 1990 this disadvantage had evaporated, and the chances of blacks, whites, and Hispanics enrolling in the college track were about equal. These findings held true regardless of whether track membership was based on student self-reports or information gleaned from school transcripts. Asian students held an advantage in college-track placement in 1990. Lucas and Gamoran's research is a reminder that achievement gaps other than black-white and Hispanic-white deserve closer attention. Because of Asian students' advantage in gaining access to high-track courses, and the achievement benefit offered there, Lucas and Gamoran conclude that race continued to be a determinant of track location in 1990 and that track location continued to maintain or exacerbate race-linked differences in achievement.

What about Washington, D.C.? Ann Flanagan and David Grissmer argue that the federal government has an important role to play in closing the achievement gap. They document that the most glaring inequalities in financial resources are between, not within, states and argue that interstate dispari-

ties necessitate a more vigorous federal effort in equalizing resources. In 1995–96, for example, after adjusting for cost-of-living differences, Mississippi spent \$4,900 per pupil while New Jersey spent \$9,090. Current federal outlays are not large enough to narrow the gap appreciably. In 1996 Mississippi received \$590 per pupil from Washington, while New Jersey received \$340 per pupil. Flanagan and Grissmer firmly believe that increased spending can make a difference in narrowing achievement differences, especially when focused on low-achieving students.

Flanagan and Grissmer, in “The Role of Federal Resources in Closing the Achievement Gap,” disaggregate NAEP scores by race, region, and locality to illustrate how achievement gaps vary geographically. The widest gaps are in the Midwest and Northeast. Students living in rural and suburban areas of the Midwest and Northeast—almost 30 percent of students nationally—score at levels rivaling the highest achieving students in the world. But students in the central cities of these regions, most of whom are impoverished and black, have some of the lowest test scores in the country. The gap is huge. In the rural South, students score lower than in other regions of the country, but the gap between white and black students is at its narrowest. Flanagan and Grissmer acknowledge that expanding the federal role in education may not always be justified. But, they argue, a larger federal role is warranted in attacking problems that transcend state boundaries. To have the greatest impact on the achievement gap, Flanagan and Grissmer recommend that federal efforts should target minimizing interstate resource differences, supporting sound educational research, and improving teacher quality.

This book is cause for encouragement. In it, some of education’s most renowned scholars present a wealth of evidence that education’s achievement gap can be bridged. They document how states, districts, and schools are making significant strides in boosting the performance of poor and minority students. Among the most promising strategies are curricula focused on reading and other core academic skills, smaller classes, getting students to take tougher courses, annual testing and reporting of disaggregated achievement data, creating schools with a culture of achievement, and vouchers offered to parents in big-city school districts. These solutions span the ideological spectrum. Both liberals and conservatives can be counted as their advocates.

Any optimism should be tempered, however, by recognizing that many obstacles remain and much work is yet to be done. The book’s success stories primarily feature programs at the elementary school level. Success stories about middle and high school levels are more rare. Sound educational programs are notoriously difficult to replicate, often breaking down and losing their effectiveness when tried under different conditions and with new stu-

dents and new teachers. Implementation problems or unintended effects may arise. The California class reduction program is a prime example, as a sudden shortage of teachers and classroom space placed school systems, especially urban systems, under great strain. Having a good idea is not enough. More experiments are needed to pinpoint the conditions under which even the most promising solutions will work optimally. Independent evaluations of costs and benefits are also needed so that public dollars go into programs getting the most bang for the buck.

Bridging the achievement gap is a national imperative. It can be done. The following pages provide a blueprint for moving beyond mere measurement of the problem and finally finding effective solutions.