

How well do we understand achievement gaps?

Eric A. Hanushek

Eric A. Hanushek is the Paul and Jean Hanna Senior Fellow at the Hoover Institution of Stanford University.

An underlying principle of U.S. social policy is that education is the key policy lever for addressing poverty. In the United States and around the world, education is almost always heavily subsidized by government. The justifications for government involvement vary, but increasingly rely on the suggestion that expanded educational investments both strengthen the national economy and improve the societal distribution of income and welfare. Education, for example, had a prominent role in the U.S. “War on Poverty,” with many of the programs developed in the 1960s continuing through today. The expansion of public colleges and universities over the past three decades has also rested on distributional arguments.

This article assesses what we currently know about the role of education in improving the welfare of the disadvantaged population by looking at one particular aspect of the subject, achievement gaps for disadvantaged students. Specifically, I review literature related to measured cognitive skills, focusing on achievement rather than school attainment. For the most part, I interpret cognitive skills as measured by student

achievement tests as a direct measure of human capital. In the end, although this is a smaller and thinner body of work than the broader topic related to education in general, I think it is fundamental to much of the rest of the literature.

An overview of achievement issues

Focusing on achievement rather than school attainment has several advantages in discussing the interaction of research and policy. First, most current policy discussions relate directly to issues of quality and student learning. For example, policy discussions of accountability and the No Child Left Behind Act of 2001 (NCLB), or of preschool and the preparation of disadvantaged students for school, are concerned with what students know at any point in time. Second, a focus on achievement allows for the fact that much of education actually takes place outside of schools. Finally, a focus on achievement allows for the possibility that other policy-relevant factors, such as health and neighborhoods, are important for education.

There are also disadvantages to focusing on achievement. With many different tests, the reliability and validity of specific measures is often unknown. Additionally, achievement may not reflect an individual’s full range of education. This disadvantage may be greater at higher levels of education;

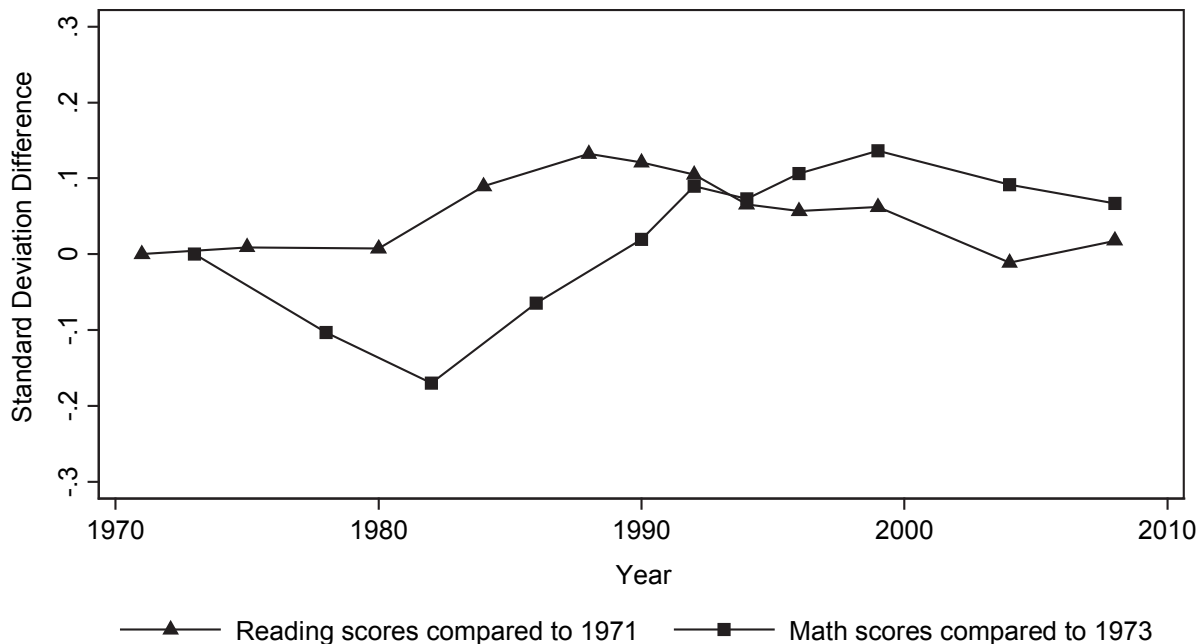


Figure 1. National Assessment of Educational Progress (NAEP) Scores, 17-year-olds 1971–2008.

Note: Figure shows standard deviation differences relative to initial performance.

Source: Author calculations based on data from <http://nces.ed.gov/nationsreportcard/>.

for example, few believe that various college tests accurately reflect what anybody has learned in college.

Until recently, school attainment has been the focus of most empirical education work, particularly as it relates to the labor market. This choice is one of convenience, because census data and other surveys have tended to measure attainment but not achievement. From the extensive evidence of inputs into educational production functions, it is apparent that school attainment is not a complete indicator of human capital and that in most situations it would need to be augmented by other determinants of achievement.¹

Returns to achievement

Three U.S. studies provide very consistent estimates of the impact of test performance on earnings of young workers.² These studies use different nationally representative datasets that follow students after they leave school and enter the labor force. When scores are standardized, they suggest that one standard deviation increase in mathematics performance at the end of high school translates into 10 percent to 15 percent higher annual earnings. Hanushek and Zhang find even larger returns to achievement (20 percent) for a more age-representative sample.³ These consistent findings demonstrate the need to pay attention to achievement gaps.

Achievement over time

The backdrop for this analysis—which emphasizes distributional, or equity, concerns—clearly must include what has been happening in terms of overall school performance. Figure 1 provides the overall performance of U.S. 17-year-olds on the National Assessment of Educational Progress (NAEP). The NAEP provides consistent national testing of a random sample of students in different subjects, so it is possible to observe any changes in performance over time.⁴ The remarkable thing about this picture is that performance appears roughly flat for almost four decades. This constancy is particularly remarkable given the amount of resources expended over that time period in an attempt to improve performance. Of the myriad changes, probably the most obvious policy response has been continued increases in the funding and resources of schools. The commonly discussed policy instruments—reduced pupil-teacher ratios, retaining more teachers, and having more-educated teachers—have been systematically employed over the past decades. Between 1960 and 2007, U.S. pupil-teacher ratios fell by 40 percent; teachers with a master’s degree more than doubled to over 50 percent; and average experience increased (see Table 1). Bringing about these changes is, of course, expensive; real spending per pupil more than tripled over the period.

The simple picture is that school policy has not been directed primarily at overall student performance (at least as seen by outcomes). Thus, it is also useful to see what happened in terms of the distribution of outcomes. This distributional discussion concentrates largely on racial differences in performance patterns because data by family income and other

Table 1
Public School Resources in the United States, 1960–2007

	1960	1980	2000	2007
Pupil-teacher ratio	25.8	18.7	16.0	15.5
Percentage of teachers with master’s degree or more	23.5 %	49.6 %	56.8 %	not available
Median years of teacher experience	11	12	14	not available
Real expenditure per student	\$3,170	\$6,244	\$10,041	\$11,674

Source: U.S. Department of Education Digest of Education Statistics, 2009 (Washington, DC: National Center for Education Statistics, 2010).

Notes: Expenditures are in 2007–2008 dollars.

measures of socioeconomic status are generally not available over time. And when such data are available, they tend to be unreliable, utilizing such measures as eligibility for free and reduced-price lunch, which is known to be incompletely reported at the high school level.

Early attention to distributional issues was provided by the massive government report, *Equality of Educational Opportunity*, commonly referred to as the “Coleman Report.”⁵ This report was mandated by the Civil Rights Act of 1964, which instructed the U.S. Office of Education to report on the lack of educational opportunity by reason of race or ethnicity. To address this issue, the Coleman research team tested some 600,000 students in the United States in 1965.

The analysis vividly underscored the huge difference in the achievement of students by race and background. A simple summary of the magnitude of differences comes from equating test scores to grade-level equivalents. If white twelfth graders in the urban Northeast (in 1965) were the standard for the knowledge that a twelfth grader should have, black students also in the urban Northeast were achieving at the ninth grade level, and black students in the rural south were achieving at the seventh grade level. The magnitude of these differences never received much attention, as most of the attention went to the analysis of the determinants of achievement.

The achievement differences have been consistent across studies. For example, when disaggregated by race, the SAT tests showed differences of approximately one standard deviation. The SAT relied on voluntary test taking for a changing group of students, however, and thus the interpretation is somewhat ambiguous. Once again, the clearest picture comes from the National Assessment of Educational Progress.

Figure 2 displays the average performance gap between whites and blacks in math and reading at age 17. Across each of the tests there is a very consistent pattern: racial gaps tended to shrink noticeably during the 1980s and then to be flat or widen somewhat during the 1990s. If anything, the white-black gap expanded in the 1990s.⁶

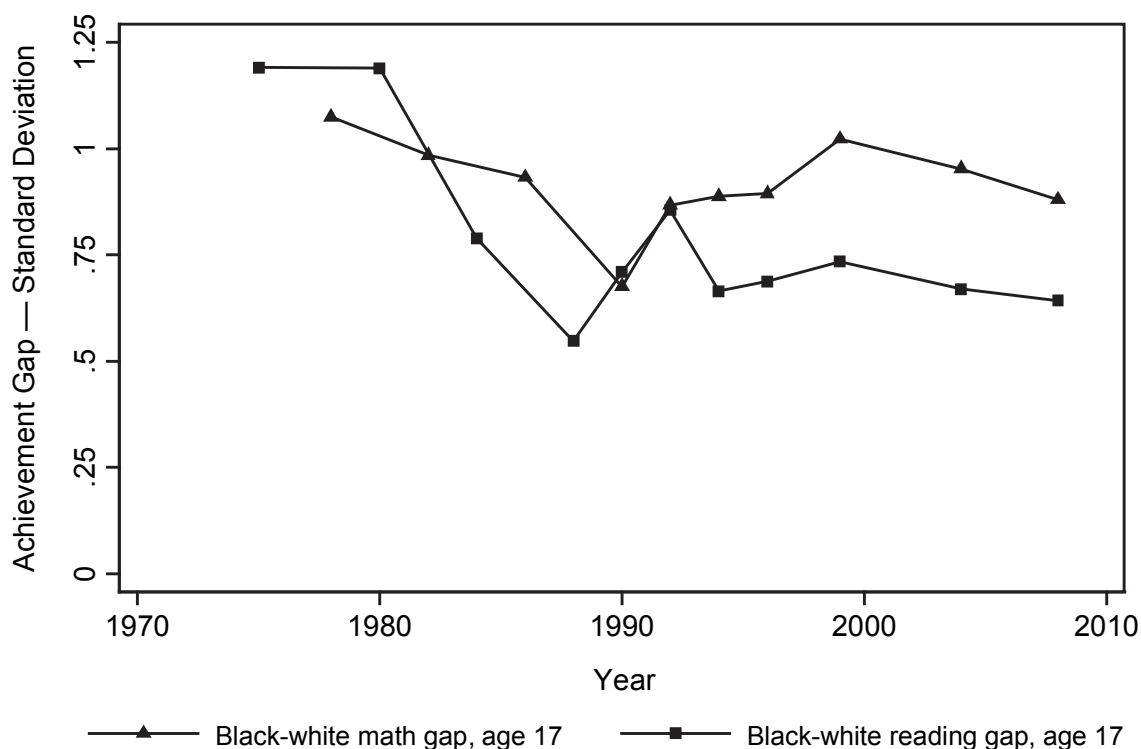


Figure 2. Black-White Achievement Gap, 1975–2008.

Source: Author calculations based on data from <http://nces.ed.gov/nationsreportcard/>.

The magnitude of the gap is stunning. The average difference in math and reading in 2008 is 0.77 standard deviations—implying that the average black 17-year-old is achieving at the 22nd percentile of the overall distribution. Other things equal, existing earnings studies indicate that the average skill differences alone imply roughly a 10 percent annual earnings difference.

Much has been made of the narrowing of the black-white achievement gap, including a widely cited conference book.⁷ The one-time nature of the test score convergence, however, was not generally anticipated and has received less attention than the significant closing of the gaps that occurred over a decade ago.

One other point of comparison is relevant. Figure 3 provides the Hispanic-white achievement gaps over time. These gaps have been flatter than the black-white gaps, and are also smaller, averaging 0.64 standard deviations across subjects in 2008. Even though there are more Hispanics than blacks in public schools (21 percent compared to 17 percent in 2007), consideration of Hispanic performance still remains limited.

Achievement outcomes

In the following section, I review studies that look at three factors that may affect the outcomes of disadvantaged students: teacher effectiveness, racial concentration in schools, and preschool education.

Distributional aspects of schools and teachers

Estimates of variations in teacher quality suggest that having a good teacher for three to five years would eliminate the average gap between children who do and do not receive free or reduced-price lunch, and between whites and blacks or Hispanics.⁸ In practice, this potential is unlikely to be realized, as few students actually get such a long exposure to good teachers.

The question remains whether or not teachers are distributed adversely for low-income and minority students. On this, the evidence is less clear. A considerable amount has been written about differences in spending across schools within large districts.⁹ These differences, however, largely reflect teacher salaries, which in turn reflect experience and graduate degrees—things mainly uncorrelated with effectiveness. Concern about uneven distribution of teachers within districts also motivated parts of NCLB that called for “highly qualified” teachers in schools serving disadvantaged students, again with little evidence that the standards were related to teacher effectiveness.

Existing evidence shows that teachers who switch schools tend to move to schools with higher achieving, higher income, and fewer minority students than their previous schools, and those changing districts tend to get slightly higher wages on average once the wages are adjusted for changes in student demographic composition.¹⁰ On average in 2005, 13 percent of teachers left schools with 5 percent

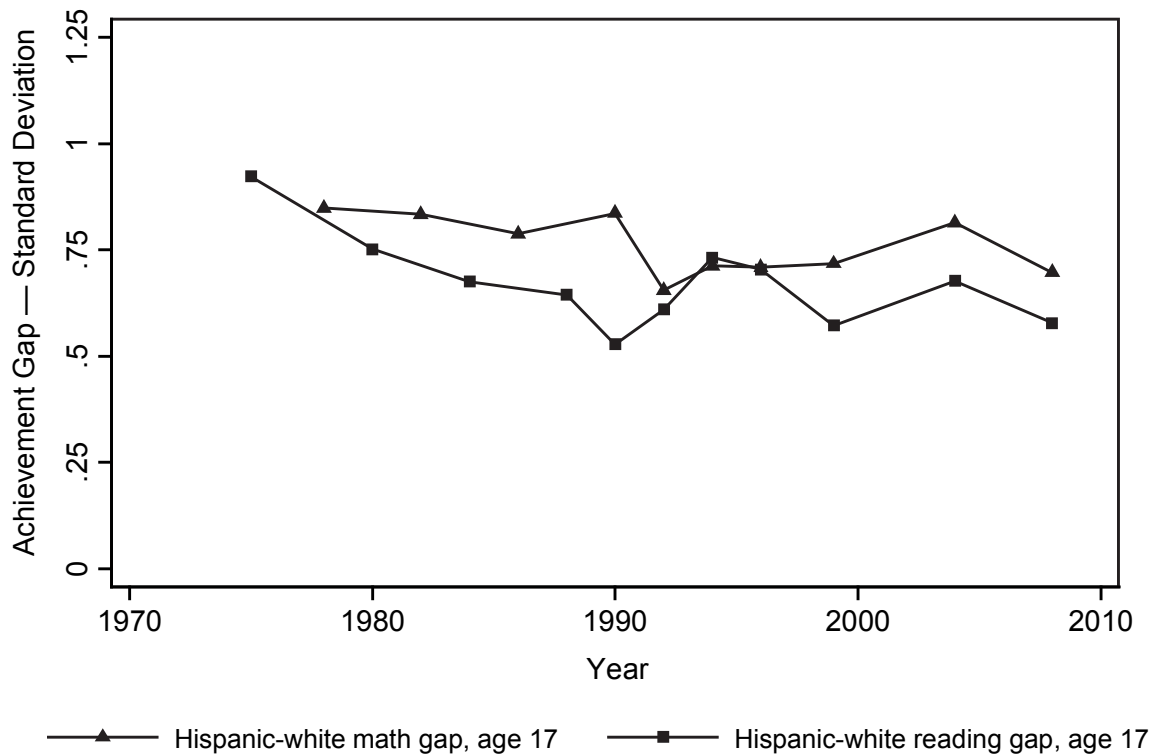


Figure 3. Hispanic-White Achievement Gap, 1975–2008.

Source: Author calculations based on data from <http://nces.ed.gov/nationsreportcard/>.

or less minority students, while 20 percent of teachers left schools with 50 percent or more minority students.¹¹ Since demographic composition is likely to be related to working conditions, these findings suggest that salary is outweighed by other considerations in job decisions of teachers. As more-experienced teachers move away, they are replaced by rookie teachers, implying that schools serving disadvantaged students will tend to have a greater proportion of new teachers. Preliminary analysis of principals finds that they follow similar mobility patterns, suggesting that administrator skills may also vary with the student population. Boyd, Lankford, Loeb, and Wyckoff also find that teacher labor markets tend to be highly localized, which further disadvantages high-poverty, lower-achieving schools located in urban centers and rural areas that tend to produce few college graduates.¹² However, even with the differences in teacher mobility, the first-year-teaching effect cannot account for much of the observed achievement gap.

Segregation and school outcomes

Poverty, race, and schooling are very highly correlated with location. A variety of people have traced different dimensions of residential locations and segregation by race and income. Cutler, Glaeser, and Vigdor describe black migration to urban areas from 1890 to 1940, which led to racial ghettos.¹³ As the migration continued between 1940 and 1970, ghettos expanded and racial segregation increased continuously. Since 1970, there has been a modest decline in segregation as blacks moved to suburban areas and central

cities became less segregated. Despite these large changes in segregation over time, segregation across cities remains very persistent and is strongly related to city size. Iceland and Weinberg examine residential segregation in metropolitan areas for the four major racial and ethnic minority groups in the United States—American Indians and Alaska Natives, Asians and Pacific Islanders, blacks, and Hispanics or Latinos.¹⁴ They conclude that blacks are the most residentially segregated of the four groups examined, but that their segregation declined between 1980 and 2000. Hispanics are the second-most-segregated group, and their overall concentrations by neighborhood have not changed over the period. Swanson, Casey, Flack, and Dreier analyzed economic segregation among municipalities for 50 major metropolitan areas.¹⁵ They conclude that economic segregation among municipalities is rising, but the trends vary significantly across time and in different regions of the country.

Fischer, Stockmayer, Stiles, and Hout study trends in residential segregation in the United States from 1960 to 2000 along several social dimensions, including race, income, and family status, and across several geographic levels: region, metropolis, the center city-suburb division, municipality, and tract.¹⁶ They report that the segregation of blacks decreased considerably after 1960, largely because neighborhoods became more integrated. While the central city-suburb barrier lessened for blacks, suburbs themselves became more segregated. The segregation of Hispanics, however, changed little. Economic segregation increased between 1970 and 1990

mainly because the affluent were clustered more in both specific metropolitan areas and in specific municipalities within metropolitan areas. An important element, however, is that economic segregation is significantly less than racial segregation.

Perhaps the most significant social policy of the 20th century was the desegregation of schools following the 1954 *Brown v. Board of Education* decision. From the late 1960s through 1980, black exposure to whites rose dramatically.¹⁷ After the late 1980s, there was some decline in white exposure, but it remained improved over the 1960s.¹⁸

Currently, the fundamental force behind school segregation is the residential location of blacks and whites across jurisdictions. In particular, completely balanced schools within districts would yield very small differences in segregation beyond that across districts.¹⁹ Moreover, minority enrollment is very much an issue in large urban areas. Over 30 percent of blacks and one-quarter of Hispanics attend schools in one of the top 50 districts.²⁰

Surprisingly, relatively little attention has been given to identifying the impacts of racial and ethnic segregation on achievement, and the available analyses provide a mixed picture. The Coleman Report provided early empirical evidence that racial isolation harms academic achievement, although Armor raises questions about the findings.²¹ Subsequent work also finds that school racial composition affects academic, social, and economic outcomes.²² On the other side, Rivkin finds no evidence that exposure to whites increases academic attainment or earnings for black men or women in the high school class of 1982; Card and Rothstein find that neighborhood but not school racial composition affects achievement; and Cook and Evans indicate that little of the black-white difference in NAEP scores can be attributed to racial concentration.²³ While varying in details, a recurring concern throughout these studies is the lack of a convincing identification strategy for uncovering the causal impact of racial concentration in the schools.

Hanushek, Kain, and Rivkin attempt to sort out the independent impact of racial composition on achievement in a framework similar to that described for teacher-effects studies.²⁴ They provide strong evidence that increases in the proportion of black students in a school adversely affects mathematics achievement of blacks. These effects are much larger and more precisely estimated for blacks than the corresponding estimated impacts on whites, which are generally not significantly different from zero. Moreover, Hispanic enrollment share exerts a far smaller effect, indicating that it is the proportion of black students rather than proportion of minority students that is the key aspect of peer race and ethnic composition in terms of achievement for blacks and whites.

The magnitudes of the black-composition effects are significant. On average, the black share of school enrollment in Texas is almost 30 percentage points higher for black

students than for white students. Elimination of this gap implies, according to the direct estimates, that the racial achievement gap would fall by 0.05 standard deviations in a single year. Such a reduction for grades 5 to 7 (the sample grade span of the estimates) suggests that a three-year cumulative effect of racial composition equalization would reduce the race achievement gap by roughly 14 percent.²⁵ Moreover, Hanushek and Rivkin suggest that it is high-achieving black students who are most harmed by increased racial concentration in schools.²⁶

Preschool education

A recent focus of policy discussions is preschool education. Various types of preschool education, such as universal or means-tested, are frequently mentioned as the next “obvious” fix for the current schooling problems, particularly for disadvantaged students who come to school far behind their middle-class peers in language and other skills.

There are three arguments for why broad provision of preschool education is a good idea. First, the problems of disadvantaged children at entry to school have received increased attention, particularly with the availability of new longitudinal data for early childhood.²⁷ The deficits in preparation of disadvantaged children are significant. For example, in evaluating the vocabulary of disadvantaged children, Hart and Risley found that they were exposed to dramatically less vocabulary. More-advantaged children at age 3 had vocabularies that were four times as large as disadvantaged 3-year-olds. Moreover, the quality of parent-child communication was vastly different. These differences in preparation have potentially lasting effects on student outcomes, where the previous charts indicate that schools have been unable, on average, to close these gaps.

Second, a variety of conceptual arguments for early investments in human capital—most notably by James Heckman and his colleagues—have received scholarly and policy attention. In a series of articles, Heckman has argued that early investments are critical, since “learning begets learning.”²⁸ Investments made early in life enhance learning later in school and even into careers, making such investments attractive.

Third, key studies with strong research designs have supported the efficacy of preschool education. The most well-known is the Perry Preschool Program, but others, such as the Abecedarian Program and the Early Training Program, also provide important evidence.²⁹ A set of benefit-cost analyses of the Perry Preschool Program shows that this appears to have been an effective program that was worth the expenditure.³⁰

Given this background, it is natural that discussions of preschool enter into the educational policy debate and into judicial proceedings and judgments. For example, courts in South Carolina and New Jersey have found preschool education to be an essential element of an adequate education.³¹

Despite the popularity of preschool programs, there are serious questions concerning the interpretation of the underlying evaluations and whether their results have general application. For example, the evaluations of the Perry Preschool, Abecedarian, and Early Training programs relied upon a random assignment methodology that followed subjects over extended periods of time, but the numbers of children taking part in the experiments were relatively small, with only around 50 children in each treatment group. Clearly, with samples of this size, one must be concerned about whether the evaluation results can be generalized to much larger programs, especially when, upon reanalysis, many of the originally reported findings have turned out to be fragile.³²

Moreover, even the beneficial results are quite varied. First, virtually all of the positive programmatic results are for females, with male children having primarily zero or negative impacts.³³ Second, a substantial part of the beneficial impact falls outside of schools and the development of cognitive skills. In particular, a substantial part of the benefits found for females relates to reduced criminal behavior.³⁴ Third, the results differ across programs, so that it is impossible simply to refer to “preschool,” but it is necessary to identify the precise kind of treatment.

Probably most important, these programs are not typical community or school-based programs found in most states. The Perry Preschool Program, estimated to cost over \$15,000 per child per year (in 2000 dollars), involved intensive treatment by teachers with master’s degrees in child development, student-teacher ratios of 6 to 1, and regular home visits.³⁵ The Abecedarian Program is full day, five days per week, 50 weeks per year, for five years beginning at birth and including medical care and intensive home visitation.³⁶ It is estimated to cost \$76,000 per child (in 2002 dollars).

Many people also forget that we have, in fact, a large public preschool program, introduced with the War on Poverty programs in 1965. Over 900,000 3- and 4-year-olds from families in poverty are currently enrolled in Head Start programs around the country. The federal Head Start program is considerably different from the Perry and Abecedarian programs. In 2005, just 35 percent of its teachers had a bachelor’s degree, and the programs varied considerably in length and intensity.³⁷ The cost of Head Start is usually reported as slightly over \$7,000 per pupil per year (in 2003–04 dollars), derived by dividing total program costs by the number of participants. However, this mixes together a variety of different programs; if run on a full-time, full-year basis, the program costs would exceed \$20,000 per pupil per year.³⁸

Support for the educational efficacy of Head Start is limited. The early education program in Head Start was complicated by its emphasis on local community employment activities, and, after initial evaluations found little lasting impact on student achievement, was redefined as a health and nutrition program instead of an educational program. Subsequent

evaluations have consistently found small achievement effects that generally disappear relatively quickly.³⁹

In 2005, 70 percent of the 4-year-olds and 5-year-olds who were not in kindergarten were in center-based care arrangements that averaged 27 hours per week.⁴⁰ Indeed, for all children ages 0 to 5, blacks (36 percent) and Hispanics (29 percent) were more likely than whites (27 percent) to be in a center-based program. (The differences largely reflected differential participation in Head Start programs.) Thus, preschool programs have already reached large portions of the young population.

In sum, there are reasons to be favorably disposed to instituting expanded preschool programs for disadvantaged students, but there are also potentially huge costs and problems associated with doing it right. The idea has been to supplement what goes on in the home in order to provide stronger educational development. Such preschool investments recognize that it is easier to remediate earlier rather than later. At the same time, the educational outcomes of existing programs that have been evaluated, except perhaps the most intensive and expensive, have been small and short lived. The limited number of models that have been evaluated provides uncertain guidance about design of effective programs, particularly programs that reach male children.

One other aspect of the design is also important. Any proposals of governmental support for preschool must consider which groups should receive programmatic help, how the programs should be organized, and how they should be financed.⁴¹ The existing evidence on preschools is limited largely to their impact on disadvantaged students. There is no evidence about positive impacts for middle- and upper-income students.⁴²

Conclusions

The starting point for this article is that achievement gaps by race (and income) are large and substantively important. Moreover, except for the gap closing during the 1980s, there has been little systematic movement.

Some conclude that schools lack the power to effect significant changes in achievement gaps. But there is a difference between having the capacity to lessen existing gaps and having an institutional structure and set of policies that accomplish such an objective.

The existing research suggests that there are three places to look for improvements. First, without a doubt, the biggest influence of schools comes through teachers. Improving teacher effectiveness could dramatically improve the achievement of disadvantaged students. Second, at least for blacks, it appears that racial concentration in schools is a significant factor. Here, the policies that would be effective are quite

unclear, given the importance of residential segregation and the force of legal restraints. Third, some sort of preschool education for disadvantaged students could potentially deal with the typical lesser preparation these students have at entry to school, yet, the exact policies and nature of any new preschool programs need to be developed. ■

¹Simple Mincer earnings models have shown that school attainment has an independent effect on individual earnings even when achievement is controlled for. However, in an international analysis, Hanushek and Zhang (E. A. Hanushek and L. Zhang, "Quality-Consistent Estimates of International Schooling and Skill Gradients," *Journal of Human Capital* 3, No. 2 (Summer 2009): 107–143) indicate that the estimated "Mincer return" falls by 40 percent on average after measures of mother's education, health, and ability are added to the Mincer earnings equation. This impact on the estimated schooling gradients is far larger than those reported in D. Card, "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, Eds. O. Ashenfelter and D. Card (Amsterdam: North-Holland, 1999): 1801–1863, where other inputs to human capital are not considered.

²See C. B. Mulligan, "Galton Versus the Human Capital Approach to Inheritance," *Journal of Political Economy* 107, No. 6, part 2 (1999): S184–S224; R. J. Murnane, J. B. Willett, Y. Duhaldorbe, and J. H. Tyler, "How important are the cognitive skills of teenagers in predicting subsequent earnings?" *Journal of Policy Analysis and Management* 19, No. 4 (Fall 2000): 547–568; and E. P. Lazear, "Teacher Incentives," *Swedish Economic Policy Review* 10, No. 3 (2003): 179–214.

³It is convenient to convert test scores into measures of the distribution of achievement across the population. E. A. Hanushek and L. Zhang, "Quality-Consistent Estimates of International Schooling and Skill Gradients," *Journal of Human Capital* 3, No. 2 (Summer 2009): 107–143, analyze data from the International Adult Learning Survey that covers a sample across all ages of adults. A separate review of the normalized impact of measured cognitive skills on earnings in S. Bowles, H. Gintis, and M. Osborne, "The Determinants of Earnings: A Behavioral Approach," *Journal of Economic Literature* 39, No. 4 (2001): 1137–1176, finds that the mean estimate is only 0.07, or slightly over half of the specific studies here.

⁴Testing is conducted at ages 9, 13, and 17, and there have been some larger changes at younger ages.

⁵J. S. Coleman, E. Q. Campbell, C. J. Hobson, J. McPartland, A. M. Mood, F. D. Weinfeld, and R. L. York, *Equality of Educational Opportunity*, Washington, DC: U.S. Government Printing Office, 1966.

⁶Scores at age 17 are obviously the product of schooling received over the prior ten years. A review of achievement gaps for 13-year-olds shows that the gains seen during the 1980s for the oldest students have their antecedents in the 1970s. Most recently, the achievement gap for 9-year-olds narrowed in reading and math. Some popular statements have attributed this narrowing to increased national accountability and particularly the introduction of the federal *No Child Left Behind Act of 2001*, but detailed analysis is unavailable.

⁷C. Jencks and M. Phillips, eds., *The Black-White Test Score Gap* (Washington, DC: Brookings, 1998).

⁸This calculation compares having an average teacher to having a teacher one standard deviation above the mean. It has uncertainties about the accumulation of gains, since it is based on one-year growth.

⁹For example, see M. Roza and P. T. Hill, "How Can Anyone Say What's Adequate if Nobody Knows How Money is Spent Now?" In *Courting Failure: How School Adequacy Lawsuits Pervert Judges' Good Intentions and Harm our Children*, ed. E. A. Hanushek (Stanford: Education Next Books, 2006), pp. 235–255.

¹⁰See H. Lankford, S. Loeb, and J. Wyckoff, "Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis," *Educational Evaluation and Policy Analysis* 24, No. 1 (Spring 2002): 37–62; E. A. Hanushek, J. F. Kain, and S. G. Rivkin, "Why Public Schools Lose Teachers," *Journal of Human*

Resources 39, No. 2 (2004): 326–354; and B. Scafidi, D. L. Sjoquist, and T. R. Stinebrickner, "Race, Poverty, and Teacher Mobility," *Economics of Education Review* 26, No. 2 (2007): 145–159. These three studies find that working conditions as measured by student demographic composition tend to be a more important determinant of school choice than salaries. Average salary changes following a move, and adjusted for working conditions, are positive but small.

¹¹For the high minority schools, those leaving teaching altogether and those changing schools were roughly the same percentages; see U.S. Department of Education, *Digest of Education Statistics*, 2008, Washington, DC: National Center for Education Statistics, 2009, Table 73.

¹²D. Boyd, H. Lankford, S. Loeb, and J. Wyckoff, "The Draw of Home: How Teachers' Preferences for Proximity Disadvantage Urban Schools," *Journal of Policy Analysis and Management* 24, No. 1 (Winter 2005): 113–132.

¹³D. M. Cutler, E. L. Glaeser, and J. L. Vigdor, "The Rise and Decline of the American Ghetto," *Journal of Political Economy* 107, No. 3 (1999): 455–506.

¹⁴J. Iceland and D. H. Weinberg, *Racial and Ethnic Residential Segregation in the United States: 1980–2000, Census 2000 Special Reports*, Washington, DC: U.S. Bureau of the Census, 2002.

¹⁵T. Swanstrom, C. Casey, R. Flack, and P. Dreier, "Pulling Apart: Economic Segregation among Suburbs and Central Cities in Major Metropolitan Areas, 1980–2000," In *Redefining Urban and Suburban America: Evidence from Census 2000*, eds. B. Katz, A. Berube, and R. Lang (Washington, DC: Brookings Institution, 2006): pp. 143–165.

¹⁶C. S. Fischer, G. Stockmayer, J. Stiles, and M. Hout, "Distinguishing the Geographic Levels and Social Dimensions of U.S. Metropolitan Segregation, 1960–2000," *Demography* 41, No. 1 (2004): 37–59.

¹⁷S. G. Rivkin and F. Welch, "Has School Desegregation Improved Academic and Economic Outcomes for Blacks?" In *Handbook of the Economics of Education*, eds. E. A. Hanushek and F. Welch (Amsterdam: North-Holland, 2006), pp. 1019–1049; and C. T. Clotfelter, *After Brown: The Rise and Retreat of School Desegregation* (Princeton, NJ: Princeton University Press, 2004).

¹⁸Hanushek (E. A. Hanushek, "Black-White Achievement Differences and Governmental Interventions," *American Economic Review* 91, No. 2 (2001): 24–28) pointed out that the only obvious factor that could explain the pattern of NAEP achievement gaps between blacks and whites was the implementation of school desegregation in the aftermath of *Brown*.

¹⁹Rivkin and Welch, "Has School Desegregation Improved Academic and Economic Outcomes for Blacks?"

²⁰U.S. Department of Education, *Digest of Education Statistics*, 2007, Washington, DC: National Center for Education Statistics, 2008.

²¹D. J. Armor, "The Evidence on Busing," *Public Interest* 28 (Summer 1972): 90–126.

²²See, for example, E. A. Hanushek, and M. E. Raymond, "Does School Accountability Lead to Improved Student Performance?" *Journal of Policy Analysis and Management* 24, No. 2 (Spring 2005): 297–327.

²³S. G. Rivkin, "School Desegregation, Academic Attainment, and Earnings," *Journal of Human Resources* 35, No. 2 (Spring 2000): 333–346; D. Card and J. Rothstein, "Racial Segregation and the Black-White Test Score Gap," *Journal of Public Economics* 91, No. 11–12 (2007): 2158–2184; M. D. Cook and W. N. Evans, "Families or Schools? Explaining the Convergence in White and Black Academic Performance," *Journal of Labor Economics* 18, No. 4 (2000): 729–754.

²⁴E. A. Hanushek, J. F. Kain, and S. G. Rivkin, "New Evidence About *Brown v. Board of Education*: The Complex Effects of School Racial Composition on Achievement," *Journal of Labor Economics* 27, No. 3 (2009): 349–383.

²⁵If the racial composition factors were similar for earlier grades, this change in racial composition would mean closing the seventh grade achievement gap by 21 percent.

²⁶E. A. Hanushek and S. G. Rivkin, “Harming the Best: How Schools Affect the Black-White Achievement Gap,” *Journal of Policy Analysis and Management* 28, No. 3 (Summer 2009): 366–393.

²⁷See the description of the three separate panels created under the Early Childhood Longitudinal Program (ELCS) at <http://nces.ed.gov/ecls>.

²⁸See, for example, F. Cunha, J. J. Heckman, L. Lochner, and D. V. Masterov, “Interpreting the Evidence on Life Cycle Skill Formation.” In *Handbook of the Economics of Education*.

²⁹A comprehensive description and evaluation of different preschool programs can be found in D. J. Besharov, P. Germanis, C. Higney, and D. M. Call, “Summaries of Twenty-Four Early Childhood Evaluations,” Welfare Reform Academy, College Park, MD: University of Maryland, 2008.

³⁰See, for example, C. R. Belfield, M. Nores, S. W. Barnett, and L. J. Schweinhart, “The High/Scope Perry Preschool Program,” *Journal of Human Resources* 41, No. 1 (Winter 2006): 162–190.

³¹*Abbeville v. South Carolina* (on appeal) and *Abbott v. Burke*; see E. A. Hanushek and A. A. Lindseth, *Schoolhouses, Courthouses, and Statehouses: Solving the Funding-Achievement Puzzle in America’s Public Schools* (Princeton, NJ: Princeton University Press, 2009).

³²An extensive re-analysis of the data from these programs has been conducted by Anderson: M. L. Anderson, “Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects,” *Journal of the American Statistical Association* 103, No. 484 (December 2008): 1481–1495. He attempts to correct for attrition and multiple outcome evaluations along with statistical innovations.

³³Anderson, “Multiple Inference and Gender Differences in the Effects of Early Intervention.”

³⁴The impact of differences in criminal activity are particularly important in the case of the benefit-cost analyses; see Gramlich, Edward M. 1986. “Evaluation of Education Projects: The Case of the Perry Preschool Program.” *Economics of Education Review* 5, No. 1: 17–24. The females did, nonetheless, generally have positive school completion results; Anderson, “Multiple Inference and Gender Differences in the Effects of Early Intervention.”

³⁵Cost estimates and programmatic comparisons are found in J. F. Witte, “A Proposal for State, Income-Targeted, Preschool Vouchers,” *Peabody Journal of Education* 82, No. 4 (2007): 617–644.

³⁶F. A. Campbell and C. T. Ramey, “Cognitive and School Outcomes for High-Risk African-American Students at Middle Adolescence: Positive Effects of Early Intervention,” *American Educational Research Journal* 32, No. 4 (Winter 1995): 743–772.

³⁷Head Start Bureau, *Biennial Report to Congress: The Status of Children in Head Start Programs*, Washington, DC: U.S. Department of Health and Human Services, 2005.

³⁸D. J. Besharov, J. A. Myers, and J. S. Morrow, “Costs Per Child for Early Childhood Education and Care: Comparing Head Start, CCDF Child Care, and Prekindergarten/Preschool Programs (2003/2004),” Welfare Reform Academy, Washington, DC: American Enterprise Institute for Public Policy Research, August 2007.

³⁹See, for example, the review in M. A. Vinovskis, “Do Federal Compensatory Education Programs Really Work? A Brief Historical Analysis of Title I and Head Start,” *American Journal of Education* 107, No. 3 (1999): 187–209.

⁴⁰U.S. Department of Education, *Digest of Education Statistics, 2009*, Washington, DC: National Center for Education Statistics, 2010, Table 44.

⁴¹For a sense of the disagreements about programs and purposes, see D. L. Kirp, *The Sandbox Investment: The Preschool Movement and Kids-First Politics* (Cambridge, MA: Harvard University Press, 2007); B. Fuller, *Standardized Childhood: The Political and Cultural Struggle Over Early Education* (Stanford, CA: Stanford University Press, 2007); and C. E. Finn Jr., *Reroute the Pre-School Juggernaut* (Stanford, CA: Hoover Institution Press, 2009).

⁴²Fuller, *Standardized Childhood*, chapter 6.