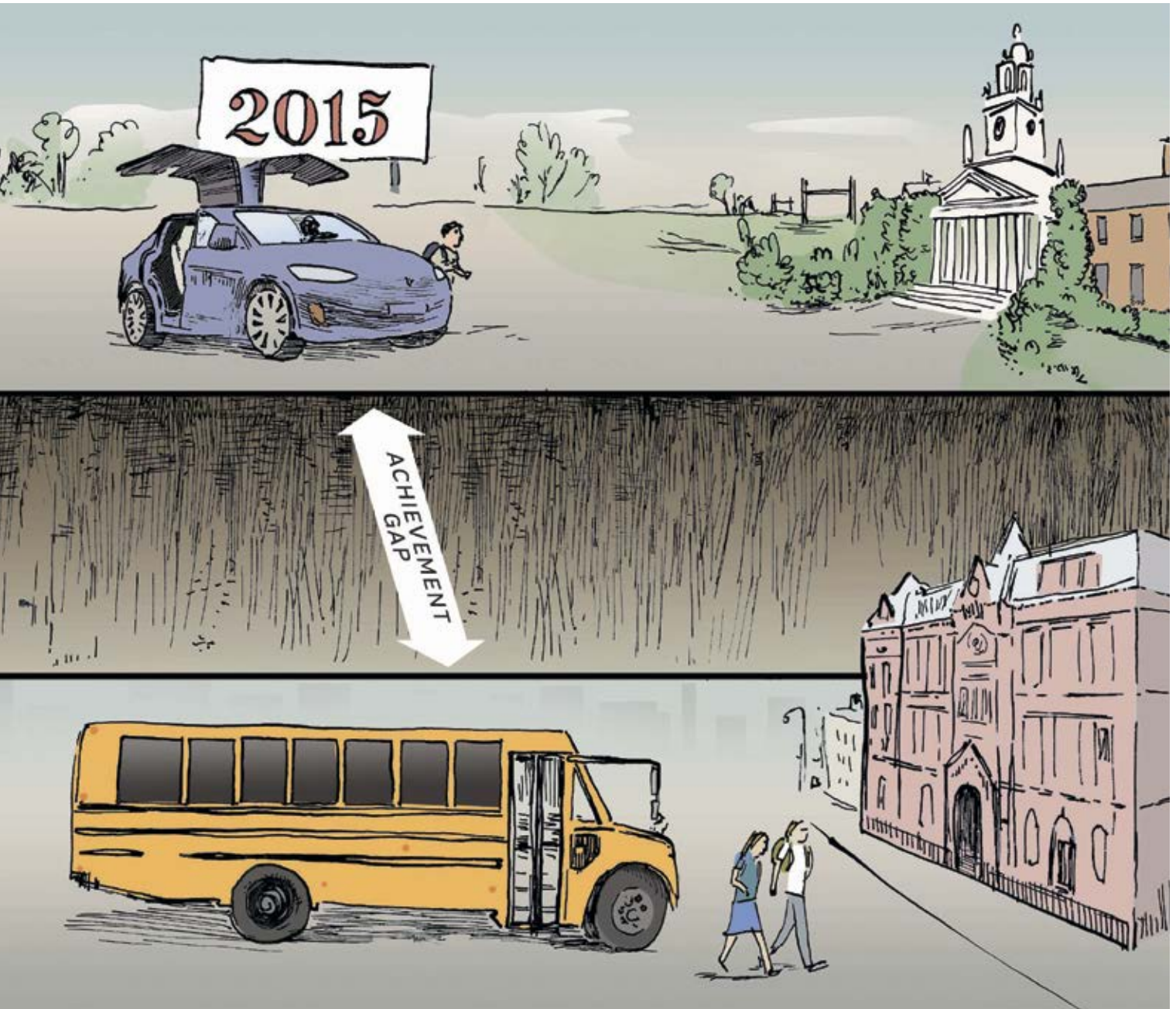ILLUSTRATION / KURT HOFFMAN

The Achievement Gap
FAILS TO CLOSE

Half century of testing shows persistent divide between haves and have-nots

INCOME INEQUALITY HAS SOARED in the United States over the past half century. Has educational inequality increased alongside, in lockstep?

Of course, say public intellectuals from across the political spectrum. As Richard Rothstein of the liberal Economic Policy Institute puts it: "Incomes have become more unequally distributed in the United States in the last generation, and this inequality contributes to the academic achievement gap." Harvard political scientist Robert Putnam, citing research by Stanford sociologist Sean Reardon, says,

"Rich Americans and poor Americans are living, learning, and raising children in increasingly separate and unequal worlds." Another well-known political scientist, Charles Murray, argues that "the United States is stuck with a large and growing lower class that is able to care for itself only sporadically and inconsistently.... The new upper class has continued to prosper as the dollar value of the talents they bring to the economy has continued to grow."

These analysts have good reason to express concern. National competitiveness is at stake, as education advocates

by ERIC A. HANUSHEK, PAUL E. PETERSON, LAURA M. TALPEY, and LUDGER WOESSMANN

have argued since the Soviet *Sputnik* launch inspired the National Defense Education Act of 1958. Economic productivity and growth are greater in countries where students perform better in math, reading, and science than in those that do not provide their youth the same opportunities to learn (see "Education and Economic Growth," *research,* Spring 2008). And while some might see income inequality as the result of adult life choices about matters such as how hard to work or where to live, educational inequality seems unfair, because the economic status of a child is outside the child's own control. It is an inequality of opportunity that runs counter to the American dream.

Despite the topic's importance, surprisingly little scholarship has focused on long-term changes in the size of the achievement gap between students from higher and lower socioeconomic backgrounds. Our new research, presented here, attempts to fill this void, using data from four national assessments of student performance administered to representative samples of U.S. students over nearly five decades.

Contrary to recent perceptions, we find the opportunity gap—that is, the relationship between socioeconomic status and achievement—has not grown over the past 50 years. But neither has it closed. Instead, the gap between the haves and have-nots has persisted.

The stubborn endurance of achievement inequalities suggests the need to reconsider policies and practices aimed at shrinking the gap. Although policymakers have repeatedly tried to break the link between students' learning and their socioeconomic background, these interventions thus far have been unable to dent the relationship between socioeconomic status and achievement. Perhaps it is time to consider alternatives.

Before drawing this conclusion, though, it is important to document the long-term trends in the connection between socioeconomic background and school achievement. Press coverage of the subject typically mentions only the most recent shifts in achievement levels and gaps. Our study broadens the perspective by making full use of nearly 50 years' worth of historical data available from four intertemporally linked assessments of achievement in math, reading, and science administered to nationally representative samples of adolescent students born between 1954 and 2001. (By "intertemporally linked," we mean that the test makers in each of these assessments design the tests to be comparable over time by doing things such as repeating some of the same questions across different waves.) These testing programs also collect information on students' socioeconomic backgrounds, which we use to construct an index of socioeconomic status. We report changes in the gaps in performance between students from more- and less-advantaged backgrounds over the past half century.

We find that the socioeconomic achievement gap among the 1950s birth cohorts is very large—about 1.0 standard deviations between those in the top and bottom deciles of the socioeconomic distribution (the "90–10 gap") and around 0.8 standard deviations between those in the top and bottom quartiles (the "75–25 gap"). These are very extensive disparities, as 1 standard deviation is approximately the difference in the average performance of students in 4th and 8th grades, or four years' worth of learning. But though these inequalities are large, they have neither increased nor decreased significantly over the past 50 years.

It could be, however, that the picture is not as dismal as suggested. If overall changes in society, coupled with policy initiatives, have proportionately lifted all boats at the same rate, everybody might be better-off, even if gaps have not significantly changed. Using the same data as for the gap analysis, we find gains in average student performance of about 0.5 standard deviations for students at age 14, or roughly 0.1 standard deviations per decade. But, surprisingly, over the last quarter century, those gains disappear for students by age 17. In other words, there is no rising tide for students as they leave school for college and careers.

Prior Research

The effects of family background on student achievement are well-documented, but few studies track changes in the relationship between demographic characteristics and student performance over time. This scarcity of longitudinal analysis partly reflects measurement challenges.

Family background and achievement. There is little dispute that students' performance in school is strongly affected by their family background. James Coleman and colleagues, in their seminal 1966 study, *Equality of Educational Opportunity,* found that parental education, income, and race are strongly associated with student achievement, while school resources such as per-pupil expenditures and class size are much less significant. Subsequent research has confirmed these early findings (see "How Family Background Influences Student Achievement," *features,* Spring 2016).

A variety of mechanisms link socioeconomic status to achievement. For instance, children growing up in poorer households and communities are at greater risk of traumatic stress and other medical problems that can affect brain development. College-educated mothers speak more frequently to their infants, use a larger vocabulary with their toddlers, and are more likely to use parenting practices that respect the autonomy of a growing child. Higher-income families have access to more-enriching schooling environments, and they generally do not face the high rates of violent crime experienced by those in extremely impoverished communities. All these and other

childhood or adolescent experiences contribute to profound socioeconomic disparities in academic achievement.

Trends in the socioeconomic achievement gap. Despite firm documentation of a strong connection between socioeconomic status and student achievement, only two studies provide information on trends in the opportunity gap over time. In an appendix table of a 1998 paper, Larry Hedges and Amy Nowell report the relationship between student performance and several background characteristics across six nationally representative surveys administered between 1965 and 1992. Among these variables, parental education has the strongest correlation with student achievement, and that connection endures over time. The correlation between achievement and family income in the six surveys is weaker and declines over time.

In a second investigation, published in 2011, Sean Reardon

> Contrary to recent perceptions, we find the opportunity gap—the relationship between socioeconomic status and achievement—has not grown over the past 50 years. But the gap between the haves and have-nots has persisted.

draws on data from 12 surveys that contain information on both student achievement and reports of parental income to estimate gaps in math and reading performance of students at the 90th and the 10th percentiles of the household income distribution. In contrast to Hedges and Nowell, he finds that the "income achievement gaps among children born in 2001 are roughly 75 percent larger than the estimated gaps among children born in the early 1940s." For those born after 1974, children in families at the median income were falling farther behind those at the 90th percentile, leading Reardon to conclude that "The 90/50 gap appears to have grown faster than the 50/10 gap during the 1970s and 1980s."

Reardon's study and its conclusions have been widely cited by both academics and in the general media, and the idea that income-related achievement gaps have dramatically increased has become contemporary conventional wisdom. In a 2012 article, the *New York Times* asserts that "while the achievement gap between white and black students has narrowed significantly over the past few decades, the gap between rich and poor students has grown substantially during the same period." Another *Times* piece quotes Reardon as saying, "The

children of the rich increasingly do better in school, relative to the children of the poor. . . . This has always been true, but is much more true now than 40 years ago."

Differences between the findings reported in the two studies may be owing to the focus of Hedges and Nowell on overall correlations between socioeconomic status and achievement, while Reardon discusses disparities between the extremes of the income distribution. They could also reflect the fact that Reardon's analysis makes use of twice as many surveys as the earlier study, including data on more recent cohorts.

We, however, explore a third possibility—methodological limitations common to both studies. Both estimate trends from data collected by different surveys that are administered to students of varying ages and use disparate methods of estimating achievement levels and socioeconomic characteristics. As Federal Reserve economist Eric Nielsen points out, when "data sources have income and achievement measures that do not map easily across surveys, they add an additional layer of complexity and uncertainty to the analysis." It is this uncertainty that we seek to mitigate by relying on surveys that allow for consistent, intertemporally linked measures of both student achievement and socioeconomic status.

Method

We draw data from four testing programs: two that are part of the National Assessment of Educational Progress (NAEP)—the Long-Term Trend and Main NAEP; the Trends in International Mathematics and Science Study (TIMSS); and the Program for International Student Assessment (PISA). (See sidebar for details.) We include all tests administered to students age 14 or thereabouts and at age 17. (For convenience, we identify all those tested at ages 13 to 15 as "14 years old.") All told, we compile observations of achievement levels and gaps from 46 tests in math, 40 in reading, and 12 in science, or a total of 98 intertemporally linked tests over a 47-year period. Across this time span, achievement data are available for 2,737,583 students.

To measure these students' socioeconomic status, we use indicators of parental education and home possessions as reported by students to construct an index similar to one designed by PISA. The choice of indicators is determined by the fact that all four assessments collect information on family background directly from students themselves. Young people are thought to be aware of their parents' level of educational attainment but to have only an imperfect knowledge of their parents' earned income. As a classic study investigating this question puts it, income is "a matter of speculation for many students and thus inaccurately reported." For this reason, the surveys collect economic information by asking students about household items, such as the number of durable goods and educational items present in the home. Students are likely

not to have seen their family's tax return. These same students, though, are well aware of whether they sleep in their own bedroom or share one. They also know whether their home includes a dishwasher or a computer. Our analysis thus differs from Reardon's study, which excludes assessments that do not ask students or their parents a direct question about household income.

We use our constructed index to estimate two disparities for each test: 1) the difference in achievement between the highest and lowest deciles of the socioeconomic distribution

(the 90–10 gap) and 2) the difference between the highest and lowest quartiles (the 75–25 gap). We then fit simple quadratic trend lines through these points in order to document how, if at all, the magnitude of these disparities has changed over time.

Achievement Gaps

As can be seen in Figure 1, the disparities in achievement between students from the highest and lowest socioeconomic status groups are strikingly persistent throughout

Data SOURCES

WE USE SURVEYS from four testing programs to investigate achievement gaps and levels over time. These surveys use consistent data-collection procedures to trace the achievement of representative samples of U.S. adolescents over time. They also collect information about the cultural and economic resources of the students' families using student reports of their parents' education and of a wide variety of durable material and educational possessions in the home. Each data set comprises student-level data that we aggregate by demographic group.

LONG-TERM TREND NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (LTT-NAEP)

The LTT-NAEP dates back to 1971 and assesses students age 9, 13, and 17. Data are available for math in select years from 1978–2008 and for reading from 1971–2008. We create a panel of math and reading scores for students age 13 and 17, beginning with the 1954 birth cohort, who turned 17 in 1971. LTT-NAEP is the only source of information for cohorts born between 1954 and 1976. In a typical year, approximately 17,000 students participate.

MAIN NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (MAIN NAEP)

The Main NAEP started in 1990 and assesses students in grades 4, 8, and 12 every two to four years. We create a panel of math and reading scores for 8th graders from 1990–2013. The Main NAEP is aligned to school curricula and designed to provide results for representative samples of students in the United States as a whole and for each participating state. For each test administration, the Main NAEP 8th-grade sample is over 150,000 observations.

TRENDS IN INTERNATIONAL MATHEMATICS AND SCIENCE STUDY (TIMSS)

TIMSS, administered by the International Association for Evaluation of Educational Achievement (IEA), is the current version of an international survey that originated as an exploratory study of mathematics conducted across 12 countries in the 1960s. The tests are curriculum-based and developed by an IEA-directed international committee. Beginning with the 1981 birth cohort (tested in 1995), the TIMSS tests have been designed to generate scores that are comparable over time. We use the TIMSS 8th-grade math and science tests beginning with this cohort by compiling national data files from 2003, 2007, and 2011, and international data files from 1995, 1999, and 2015. The only difference between the national and international data is that the latter do not contain an indicator of race or ethnicity. For this reason, our estimates of the black-white achievement gap for TIMSS are only available for 2003, 2007, and 2011. The U.S. TIMSS 8th-grade sample includes roughly 10,000 students for each administration of the test.

PROGRAM FOR INTERNATIONAL STUDENT ASSESSMENT (PISA)

PISA, administered by the Organization for Economic Co-operation and Development, began in 2000 and assesses students' math, reading, and science literacy at age 15 every three years. Its assessments are designed to measure practical applications of knowledge. The United States has participated in every wave of the test, though results are not available for reading for the 1991 birth cohort. We use national PISA data, available every three years from 2000 to 2015. PISA does not collect information on race or ethnicity, so these tests are not used in our analysis of the black-white achievement gap. The U.S. PISA sample includes over 5,000 students for each administration of the test.

A Persistent Achievement Gap between Haves and Have Nots (Figure 1)

For students born between 1954 and 2001, the achievement gap between students in the top and bottom deciles of the socioeconomic distribution remained as wide as a full standard deviation—the equivalent of three to four years of learning. The gap between students in the top and bottom quartiles also persisted over this near half century.

Differences in test performance by socioeconomic status
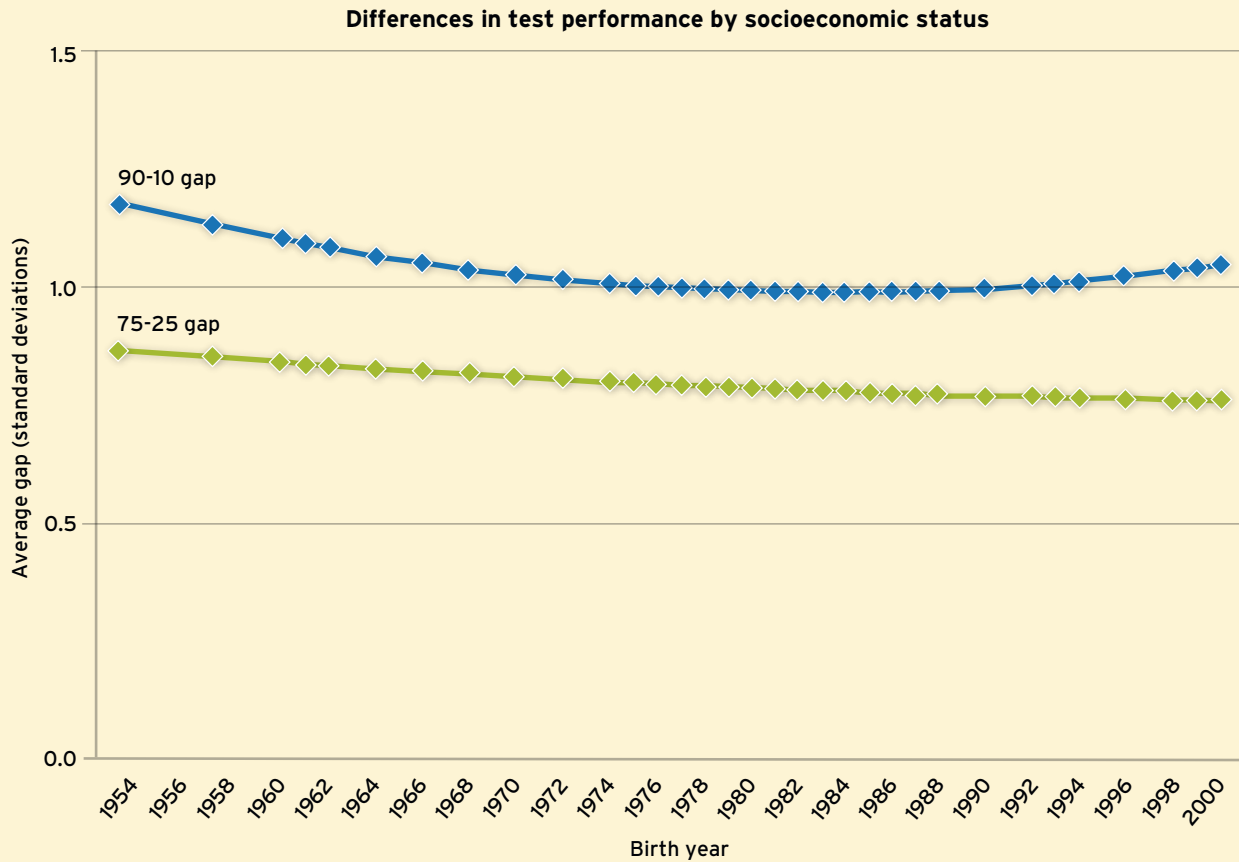
NOTE: Markers indicate years a test was administered.

SOURCE: Tests administered by LTT-NAEP, Main-NAEP, PISA, and TIMSS

the time period. The socioeconomic achievement divide hardly wavers over this half century. In the 1954 birth cohort, the achievement gap between the average of those in the top and bottom deciles of the socioeconomic distribution stood at slightly less than 1.2 standard deviations. For those born in 2001, the gap is only slightly less—about 1.05 standard deviations. That is, the most-disadvantaged students have made the same gains in achievement over the decades as those realized by the most-advantaged students.

The disparity between students in the top and bottom quartiles of the socioeconomic distribution was about 0.9 standard deviations for the 1954 birth cohort. This 75–25

gap falls slightly during the next two decades, settling at barely below 0.8 for the cohort born in 2001.

Trends are similar for math and reading separately. The gap in math achievement, particularly for the 90–10 comparison, shows a little movement over the period—narrowing in the early years but returning to a position below the initial level in recent decades. The 75–25 math gap narrows slightly over time. In reading, the pattern appears essentially flat for the entire period.

To see whether an alternative measure of socioeconomic status yields similar results, we estimate the gap between students who are eligible for the federal school-lunch

program and those who are not, as reported on the Main NAEP, the one assessment that contains this information. The federal program provides free lunch to extremely poor students from households below the poverty line, while a reduced-price lunch is available to moderately poor students with somewhat higher incomes (1.85 times the poverty line). The gap between the extremely poor students and other students in the 1982 birth cohort is a sizable 0.73 standard deviations (Figure 2). When the extremely poor are combined with the moderately poor, the gap for this cohort is nearly as large. Over the next 20 years, the gap between the extremely poor and students from families above the eligibility line narrows by just 0.02 standard deviations, while the gap between ineligible students and all those eligible for participation in the program widens by 0.01. In sum, this alternative measure of the achievement gap between students from higher and lower socioeconomic backgrounds also shows only minuscule change over the course of the past two decades.

Figure 2 also shows the white-black achievement gap. While this is not accurately thought of as a socioeconomic gap because of the improvements in black incomes, it represents another potential dimension of continuing societal disparities. As Figure 2 shows, there is a sizable shrinking of the racial gap

Our results confirm Reardon's finding of large gaps in academic performance between students at the extremes of the socioeconomic distribution . . . but not his finding that achievement differentials have risen by as much as 75 percent over the past 50 years.

in the early period but little change across the last two decades.

Some have hypothesized that the lack of success in diminishing the size of the socioeconomic gap is due to changes in the racial and ethnic composition of the school population. It is true that the ethnic makeup of the school-age population has changed dramatically over the past half century, with the share that is white declining from about 75 percent to 55 percent. However, these changes do not seem to have materially affected trends in performance gaps. The 90–10 socioeconomic achievement gap among white students born in 1954 was one standard deviation. By the middle of the

period, the divide had declined by about 0.2 standard deviations, but it then rose again by a commensurate amount. Trends for the 75–25 socioeconomic achievement gap among whites are much the same, confirming that changes in the ethnic composition of student cohorts do not account for the unwavering divide between the haves and have-nots.

In sum, our results confirm Reardon's finding of large gaps in academic performance between students at the extremes of the socioeconomic distribution. The average 90–10 income achievement gap across the surveys suggested by the Reardon analysis is very similar to the 90–10 socioeconomic achievement gap we identify. We are, however, unable to replicate Reardon's finding that achievement differentials have risen by as much as 75 percent over the past 50 years. His results may be a function of a reliance upon cross-sectional studies that use disparate methods for collecting both income and achievement information. Whatever the reason, the trends estimated in his analysis differ markedly from the gaps we observe by using a uniform measure of socioeconomic status and data from intertemporally linked surveys administered to students of the same age.

Rising Tides?

We might feel differently about these persistent achievement gaps if we found that all achievement was rising and thus suggesting improved economic futures for all. To place the achievement gaps in context, we describe changes in the average level of achievement among students at age 14 and age 17 for students born between 1954 and 2001. Figure 3 shows a significant upward trend in the average achievement level for all adolescent students of approximately 0.3 standard deviations over the course of the past half century, or approximately 0.06 per decade. This trend differs by the age of the student, however. Students at age 14 show an overall increase of about 0.43 standard deviations, or approximately 0.08 per decade, but gains among students at age 17 amount to only about 0.10 standard deviations, or 0.02 per decade. Further, we see no improvement in the performance of older students after the 1970 birth cohort.

Trends in average levels of achievement do differ in magnitude by subject, but the overall patterns are quite similar. In math, the younger adolescents register average gains of 0.9 standard deviations, while the older ones show a shift upward of only 0.25. At both ages, the reading gains are less. The trend among younger adolescents amounts to just 0.20 standard deviations over the half century and, among older ones, the trend is flat, showing no upward trend at all.

The differences in trend lines for students at different ages presents a puzzle for which we have no easy answer. Even setting aside the oldest students in our data, we see that the average improvement in test performance among

Stubborn Race and Socioeconomic Gaps (Figure 2)

The racial achievement gap shrank by roughly two years' worth of learning (from 1.3 to 0.8 standard deviations) between cohorts born in 1954 and the 1980s, but the remaining gap has persisted in recent decades. The gap between poor students eligible for free and reduced-price lunch and other students barely changed for cohorts born over the past two decades.

Achievement gaps by race and subsidized lunch eligibility
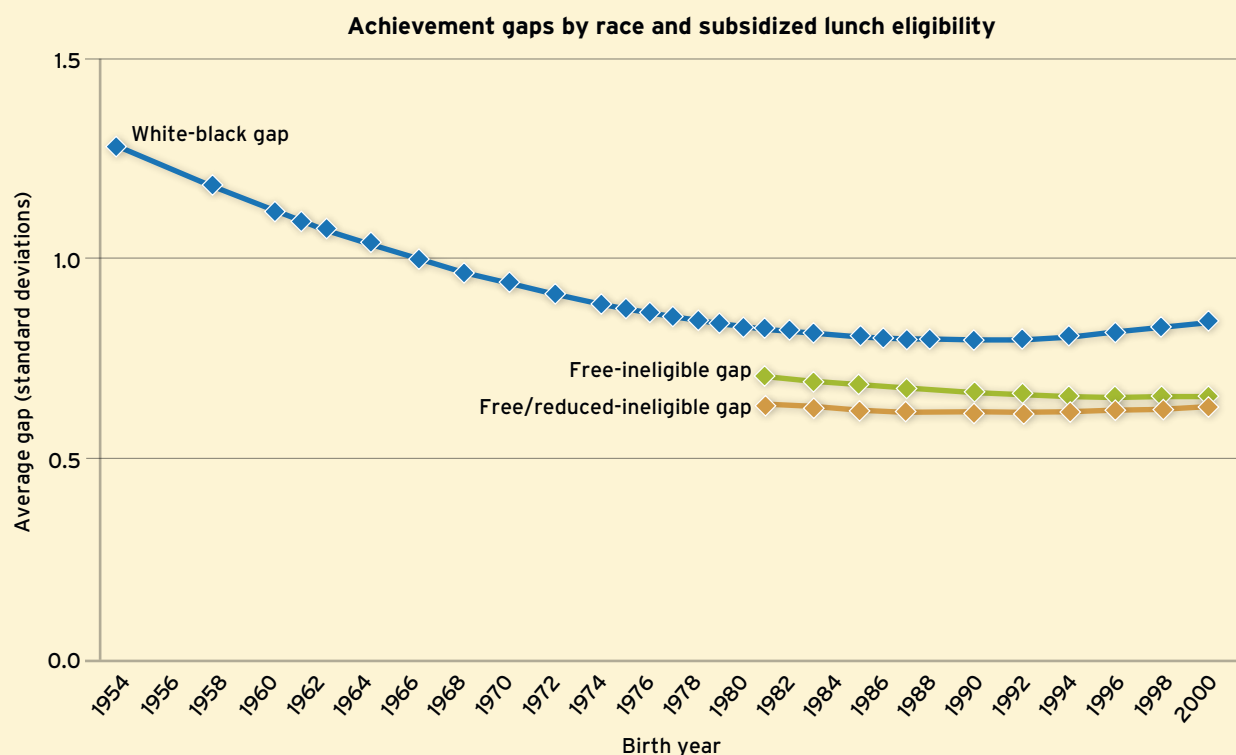
NOTE: Data on free and reduced-price lunch eligibility are only available for Main-NAEP tests, starting with the 1982 birth cohort. Markers indicate years a test was administered.

SOURCE: LTT-NAEP and Main-NAEP surveys

13- and 14-year-olds who take the NAEP tests and the TIMSS is larger than that registered by 15-year-olds on the PISA tests. This may reflect differences in test design, or it may suggest that the fade-out in gains begins in the early years of high school. The lack of a positive trend among 17-year-olds for the past quarter century also suggests that high schools do not build upon gains achieved earlier, a signal, perhaps, that the high school has become a troubled institution. In any event, there is no sign of a rising tide that lifts all boats at age 17 when these students are going into further schooling or into the labor force.

Importantly, the age anomaly that we see in the trends in achievement levels is not found in the performance gaps. Constant social gaps are found across all age groups.

Discussion

The achievement gap between haves and have-nots in the U.S. remains as large as it was in 1966, when James Coleman wrote his landmark report and the nation launched a "war on poverty" that made compensatory education its centerpiece. That gap has not widened, as some have suggested. But neither has it closed.

The question remains: why has the gap remained constant? The tempting answer is that nothing significant enough has happened to alter its size. But this would ignore a wide variety of factors that have shifted over the years. It is more likely that some changes within families and within schools have worked to close the socioeconomic achievement gap while other changes have widened it, with

these factors largely offsetting one another.

Families. In terms of family background, there is the widening differential in household income that motivated Reardon's work. Socioeconomic differences in the age of the mother at the birth of the child have also increased in the past 50 years. The incidence of single-parent households has increased and is likewise concentrated at the lower end of the socioeconomic spectrum. Each would tend to exacerbate socioeconomic achievement gaps.

But these negative factors could be offset by other, countervailing demographic changes. Most importantly, differences among children in their parents' level of educational

Steady Gains for Younger, but not Older, Adolescents (Figure 3)

Average achievement for adolescent students increased by more than a year's worth of learning (over 0.3 standard deviations) from the 1954 to 2001 birth cohorts. However, the gains of 0.4 standard deviations realized by age 14 do not persist until age 17 in the most recent quarter century.

Achievement levels of younger and older students
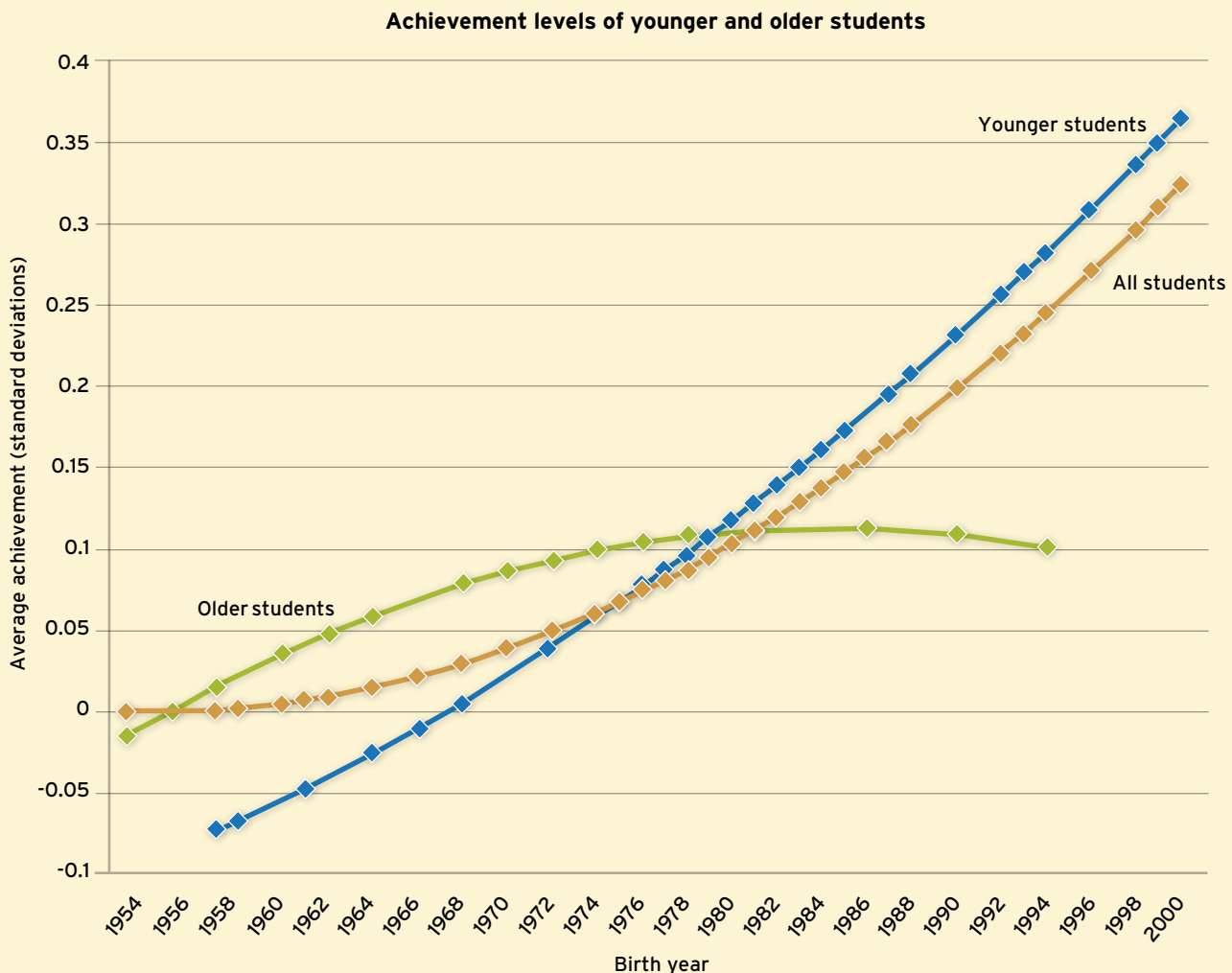
NOTE: Younger students are those between ages 13 and 15 or in 8th grade, depending on the test; the modal age is 14. Older students are those age 17 or in 12th grade, depending on the test. Markers indicate years a test was administered.

SOURCE: Tests administered by LTT-NAEP, Main-NAEP, PISA, and TIMSS

attainment have narrowed as overall education levels have climbed. So have differences in the number of siblings in the household. Both factors are important determinants of student achievement. The balance among all these factors may well have left the family contribution to the achievement gap at much the same level today as it was for cohorts born in the 1950s.

Schools. Similarly, there may be opposing forces within the educational system that have offset one another. On the one side, over the past 50 years, the federal government has

> The lack of a positive trend in achievement among 17-year-olds over the past quarter century suggests that high schools do not build upon gains achieved earlier, a signal, perhaps, that the high school has become a troubled institution over the past 50 years.

enacted compensatory education programs for school-age children and the Head Start program for students at ages three and four. *Brown v. Board of Education* and the Civil Rights Act of 1964 accelerated school desegregation, particularly in the South. The Individuals with Disabilities Education Act funded school services for students with disabilities, a group disproportionately composed of children from low-income families. States systematically changed their funding of local schools, often in response to court orders, leading to more equal funding between rich and poor school districts. Overall school funding increased dramatically on a per-student basis, quadrupling in real dollars between 1960 and 2015. And finally, states have introduced measures holding schools accountable for student performance, as required by the 2002 No Child Left Behind Act. Accountability mandates were disproportionately directed toward schools serving low-income students. Each is aimed at closing gaps.

On the other hand, the quality of the teaching force—a centrally important factor affecting student achievement—may well have declined over the course of the past several decades. Women have greater access to opportunities outside the field of teaching. Teachers' performance on standardized tests has slipped, along with other indicators of selectivity. Teacher salaries have declined relative to those earned by other four-year college-degree holders and are currently low relative to

comparable workers in other occupations (see "Do Smarter Teachers Make Smarter Students?" *features,* Spring 2019).

These changes affecting the quality of the teaching force are likely to have had a disproportionately adverse effect on disadvantaged students. Collective-bargaining agreements and state laws have granted more-experienced teachers seniority rights, leaving disadvantaged students to be taught by less-effective novices.

In other words, a growing disparity in teacher quality across the social divide may have offset the impacts of policies designed to work in the opposite direction.

Conclusion

Two surprises emerge from this analysis of long-term trends in student-achievement levels and gaps across the socioeconomic distribution. First, gaps in achievement between the haves and have-nots are mostly unchanged over the past half century. Second, steady gains in student achievement at the 8th-grade level have not translated into gains at the end of high school.

Because cognitive skills as measured by standardized achievement tests are a strong predictor of future income and economic well-being, the unwavering achievement gap across the socioeconomic spectrum sends a discouraging signal about the possibilities of improved intergenerational social mobility. Perhaps more disturbing, programs to improve the education of disadvantaged students, while perhaps offsetting a decline in the quality of teachers serving such students, have done little to close achievement gaps. These steadfast disparities suggest the need to reconsider the current direction of national education policy.

Two areas for further exploration seem especially critical. First, researchers have uniformly found that teacher effectiveness is a predominant factor affecting school quality. While there has been ample commentary on teacher recruitment and compensation policies, few programs and policies at scale have directly focused on enhancing teacher quality, particularly for disadvantaged students. Second, the achievement gains realized by students at age 14 fade away by age 17, yet policymakers have left high schools—like the achievement gap itself—in many ways untouched.

Eric A. Hanushek is the Paul and Jean Hanna Senior Fellow at the Hoover Institution of Stanford University. Paul E. Peterson is senior editor of Education Next; professor of government and director of Harvard's Program on Education Policy and Governance; and a senior fellow at the Hoover Institution. Laura M. Talpey is a research associate at the Stanford Institute for Economic Policy Research. Ludger Woessmann is professor of economics at the University of Munich.