

DRAFT April 2014: Forthcoming in *Handbook of Research in Education Finance and Policy*, Helen A. Ladd, Margaret E. Goertz (Eds.), Lawrence Erlbaum.

Patterns and Trends in Racial/Ethnic and Socioeconomic Academic Achievement Gaps

Sean F. Reardon, Joseph P. Robinson-Cimpian, and Ericka S. Weathers

Introduction

Racial, ethnic, and socioeconomic disparities in academic achievement remain a stubborn feature of U.S. schooling. National studies consistently show that the average non-Hispanic black student scores well below the average non-Hispanic white student on standardized tests of math and reading skills, as does the average Hispanic student. Likewise, the average student from a low-income family scores much lower on such tests than students from higher-income families. Considerable attention has been focused on achievement gaps, particularly the black-white achievement gap. Scholars and educators have suggested a number of possible explanations for the gaps, and policymakers, principals, and teachers have tried a range of remedies. As this chapter documents, however, the gaps persist despite these efforts. Moreover, our understanding of the causes and patterns of these achievement gaps is far from complete.

How Have Achievement Gaps Changed Over Time?

The best source of data on how achievement gaps have changed over time comes from the National Assessment of Educational Progress (NAEP).¹ NAEP includes two different assessments of the math and reading skills of nationally representative samples of students.² The first of these—Long-Term Trend NAEP (NAEP-LTT)—is given every two to four years to a nationally

¹ For more information on NAEP, see <http://nces.ed.gov/nationsreportcard/about/>.

² Main NAEP also includes assessments in writing, science, and other subjects. We focus in this chapter on the math and reading NAEP assessments.

representative sample of children aged 9, 13, and 17. Because the tests used for NAEP-LTT have remained essentially unchanged since their first administration in the early 1970s, they provide a consistent instrument to evaluate achievement trends (see for example, Grissmer, et al., 1998; Hedges & Nowell, 1998; Neal, 2005).

The second of the NAEP assessments has been administered roughly every two years since 1990 and is sometimes referred to as “Main NAEP.” The content of the Main NAEP tests is updated periodically to reflect changing curricula, so that it is more appropriate than NAEP-LTT for providing information on how students perform on tests of the material taught in their schools. While this feature has obvious appeal for investigating current achievement patterns, the changing nature of the assessments may complicate comparisons of achievement levels across years. Trends in achievement gaps, however, are less sensitive to small changes in test content than are trends in achievement levels, so Main NAEP is also useful for assessing trends in achievement gaps. Using either test, researchers typically report gaps in standard deviation units.

One additional difference between the two NAEP datasets is worth noting. NAEP-LTT is administered to age-cohorts (9-, 13-, and 17-year-olds), while Main NAEP is administered to grade-cohorts (4th-, 8th-, and 12th-graders). Because the populations sampled for the two assessments differ from one another, and because patterns of grade retention differ among racial, ethnic, and socioeconomic groups and may differ over time, the gaps and their trends measured by NAEP-LTT and Main NAEP would likely differ from one another even if the same tests were used in each. This will complicate interpretation of trends in measured gaps, since the average grade level of each age-cohort and the age composition of each grade-cohort may differ across population groups and over time.

In Tables 1 and 2, we report trends in NAEP achievement gaps. Table 1 reports 1971-2012 trends in black-white and Hispanic-white math and reading gaps, and gaps between students

whose parents have only a high school diploma and those with at least a 4-year college degree, as measured in standard deviation units on the NAEP-LTT tests. Table 2 reports 1990-2013 trends in black-white, Hispanic-white, Mexican-white, and Asian-white math and reading gaps, measured in standard deviation units on the Main NAEP tests.³

Trends in Black-White Achievement Gaps

Trend data on black-white achievement gaps are relatively clear: the black-white achievement gaps in both math and reading narrowed from the early 1970s through the late 1980s (Grissmer, et al., 1998; Hedges & Nowell, 1998, 1999; Neal, 2005).⁴ This pattern is evident not only in NAEP, but also in SAT score trends (Ferguson, 1998) and in comparing nationally representative samples of students in other large scale studies from the 1960s through 1992 (Hedges & Nowell, 1999). The black-white gap widened in the early 1990s (Ferguson, 1998; Neal, 2005), but NAEP-

³ We report the NAEP estimates of achievement gaps in standard deviation units, using the population standard deviation for each specific grade-subject-age/year combination. There are, of course, a number of different metrics in which achievement gaps can be measured, including differences in average scale scores, differences in pooled standard deviation units, and differences in average ranks, the latter of which are the basis for a set of so-called “metric-free” gap measures (Ho & Haertel, 2006). The empirical literature on achievement gaps has generally relied primarily on standardized gap measures (see, for example, Clotfelter, Ladd, & Vigdor, 2006; Fryer & Levitt, 2004, 2005; Grissmer, Flanagan, & Williamson, 1998; Hedges & Nowell, 1999; Neal, 2005; Phillips, Brooks-Gunn, Duncan, Klebanov, & Crane, 1998; Reardon & Galindo, 2006), while the “metric-free” measures have been less widely used (for recent examples of their use, see Ho & Haertel, 2006; Neal, 2005; Reardon, 2007; Robinson & Lubienski, 2011). In addition, some of the literature on achievement gaps has relied on (un-standardized) mean differences in test scores (Hanushek & Rivkin, 2006; LoGerfo, Nichols, & Reardon, 2006; Murnane, Willett, Bub, & McCartney, 2006).

Although inferences about achievement gaps can be dependent on the choice of a test metric (see, for example, Bond & Lang, 2013; Reardon, 2007; Selzer, Frank, & Bryk, 1994), the implications of using different measures of gaps have not been thoroughly investigated. In one example comparing standardized gaps and metric-free measures, Reardon (2007) shows that metric-free measures yield patterns similar to those using standardized gap measures, while mean difference measures using different versions of the test score yield very different patterns. We use standardized gaps here for ease of interpretation and comparability with most prior research.

⁴ Although the trend is clear, different studies report the narrowing of the gap in different metrics. Neal (2006) calculates the black-white achievement gap at each assessment wave in standard deviation units, and then compares changes in these gaps across waves (see also Hedges & Nowell, 1999). Grissmer, Flanagan, & Williamson (1998), however, calculate the achievement gaps by standardizing scores relative to the standard deviation of each group’s scores at the first assessment. Although these approaches yield roughly similar trends, we think the Neal (2006) approach provides a more meaningful metric for comparing gaps across cohorts.

LTT results show that the black-white gap has been narrowing consistently since 1999 (see Table 1). Likewise, the Main NAEP results also indicate a narrowing of the black-white math and reading gaps in 4th and 8th grade since the early 1990s (see Table 2).

A wide variety of explanations have been proposed for the trends in the black-white test score gaps, though none are conclusive. Grissmer, Flanagan, & Williamson (1998) argue that trends in black and white students' family characteristics can only account for a small portion of the narrowing of the black-white gap through the late 1980s; they argue instead that school policies (including desegregation) account for much of the reduction in the gap. Among other things, they demonstrate that achievement gains by region were greatest for blacks in the southeast, which experienced desegregation, and lowest in the northeast, which experienced increased segregation. Other research also suggests that segregation patterns and trends are strongly associated with racial disparities in educational outcomes, including both achievement and attainment (Card & Rothstein, 2006; Guryan, 2004; Johnson, 2011; Vigdor & Ludwig, 2007). Most recently, Condrón, Tope, Steidl, & Freeman (2013), using data from the early 1990s through the late 2000s, found that gap trends were correlated with trends in racial school segregation, net of other time varying factors. Nonetheless, much of the research linking segregation to achievement gaps is either based on data from the 1960s and 1970s—when desegregation efforts led to large exogenous changes in segregation patterns—or is correlational in nature, which leaves some uncertainty as to whether segregation patterns are causally related to achievement gaps.

Trends in Hispanic-White Achievement Gaps

Despite the availability of NAEP-LTT data on a sizable Hispanic student population, there is relatively little research on patterns and trends of Hispanic-white achievement gaps. Reardon, Valentino, Kalogrides, Shores, & Greenberg (2013) show that there have been modest decreases in the Hispanic-white achievement gap during the last 20 years. Tables 1 and 2 present standardized

achievement gaps calculated from NAEP data. The Hispanic-white reading and math gaps trends are similar to the black-white gap trends: the gaps decreased from the mid-to-late 1970s to the mid-to-late 1980s. Then increased through 1994 and have generally been declining since 1999.⁵ The Hispanic-white reading gap is generally similar in magnitude to the black-white reading gap. In math, however, the Hispanic-white gap has been consistently smaller than the black-white gap.

Data from Main NAEP also allow for the examination of achievement patterns of Mexican-origin students relative to non-Hispanic white students (Table 2). In general, the Mexican-white gaps are very similar to the overall Hispanic-white gaps (not surprisingly, since Mexican-origin students make up roughly two-thirds of Hispanic students).

Trends in Socioeconomic Achievement Gaps

Until recently, relatively little research has explicitly examined trends in what we might call the socioeconomic status achievement gap—the gap between children from lower- and higher-income or socioeconomic status families. Reardon (2011), however, estimated the income achievement gap (which he defined as the average test score gap between children from families at the 90th percentile of the family income distribution and children from families at the 10th percentile) from a set of nationally representative samples of students born from the early 1940s to 2001. He found that the income achievement gap has grown roughly forty percent over the last few decades. These trends, he argues, likely result from the combination of rising income inequality and changes in patterns of families' investments in their children since the 1970s, though more research is needed to fully understand these trends. More recent research suggests that the income achievement gap in school readiness may have started to decline in the last decade (Reardon &

⁵ When comparing Hispanic-white achievement gaps across cohorts or grades, some caution is required. Given the dramatic increase of the Hispanic student population over the past few decades, the composition of Hispanic students tested in different grades and different years may vary greatly, and so any comparison across cohorts or grades may be confounded with the changing composition of the Hispanic student population.

Portilla, 2014), though it is unclear if this represents a reversal of the trend Reardon (2011) described or simply a short-term anomaly.

In Table 1, we make an effort to examine socioeconomic achievement gaps using parental education level as a crude indicator of socioeconomic status. Specifically, we use NAEP-LTT data to compare the math and reading achievement of students whose parents have only a high school diploma to those with at least a (four-year) college degree.

The high school-college degree gaps in math were relatively stable between 1978 and 2008, ranging from roughly a half to two-thirds of a standard deviation at age 13 and ranging from three-fifths to three-quarters of a standard deviation at age 17. In 2012, however, these gaps were higher than they have ever been in over thirty years. In reading, the parental education gaps have been generally smaller than those in math and show a little more of a trend over time. The reading gaps narrowed through the 1980s before widening again, so that the most recent estimates of the gap (roughly two-thirds a standard deviation at ages 13 and 17) are now larger than they have ever been in the previous three decades.

How Do Achievement Gaps Change as Children Progress Through School?

Both racial and socioeconomic achievement gaps emerge prior to the start of kindergarten. For more information about the pre-kindergarten development of these gaps, see the chapter by Loeb & Bassok in this volume. Research findings on the development of achievement gaps as children progress through school come from two types of studies—those using longitudinal panel data on one or more cohorts of students,⁶ and those relying on repeated cross-sectional data of the

⁶ Examples of such studies include those using panel data from nationally representative samples—such as the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K) (see www.nces.ed.gov/ecls), the National Education Longitudinal Study (NELS) (see www.nces.ed.gov/surveys/nels88), Prospects: The Congressionally Mandated Study of Educational Growth and Opportunity, and High School and Beyond (HSB) (see www.nces.ed.gov/surveys/hsb)—and those drawn from state administrative data sources in states like North Carolina, Texas, or Florida, each of which has administrative data systems allowing tracking of individual student test scores over multiple years (Clotfelter, et al., 2006; Hanushek & Rivkin, 2006).

same cohort to infer developmental patterns (such as the NAEP studies). Under the respective assumptions that attrition (from longitudinal studies) is random and that changes in cohort composition are unrelated to achievement patterns (in repeated cross-sectional studies), both types of studies will provide unbiased estimates of the development of achievement gaps as students age.⁷ Regardless of whether longitudinal or repeated cross-sectional data are used to examine the development of gaps as students progress through school, the value of comparison of the magnitude of the gaps at different ages depends on the comparability of the test metrics used at each age. In fact, different test metrics lead to dramatically different conclusions regarding how achievement gaps change with age (compare, for example, Fryer & Levitt, 2004; Fryer & Levitt, 2005; Hanushek & Rivkin, 2006; see also Reardon, 2007). Gaps measured in pooled within-age standard deviation units are both more available and likely less sensitive to violations of the interval-scaled metric assumption than other approaches, so we focus here on studies that measure gaps in this type of metric.

The Development of Black-White Achievement Gaps

Almost all research on the topic concludes that the black-white achievement gaps grow during the school years, particularly in elementary school. The most commonly-cited (and probably the best) contemporary evidence on the development of the black-white gaps in elementary school comes from ECLS-K, which includes kindergarten through eighth-grade assessment data on a nationally representative sample of students who were enrolled in kindergarten in the fall of 1998 (see Table 3). ECLS-K data show that the black-white gaps in both math and reading are sizeable at the start of kindergarten—about three-quarters and one-half of a

⁷ The assumption of no cohort change may be particularly problematic, however. For example, a sample of 9-year-old students drawn in 1991 and a sample of 17-year-old students drawn in 1999 may not represent exactly the same cohort population, since the 1999 sample would include some students not in the 1991 population (e.g., those born in 1982 who immigrated to the U.S. between 1991 and 1999), and the 1991 sample would include some students not in the 1999 population (e.g., those who dropped out of school between age 9 and age 17).

standard deviation, respectively (Fryer & Levitt, 2004; Reardon, 2007; Reardon & Galindo, 2006). Measured in standard deviation units, these gaps widen slightly through kindergarten and first grade, and then widen more rapidly between first and fifth grade, and then very slowly or not at all by eighth grade, by which time the math and reading gaps are about one full standard deviation (Benson & Borman, 2010; Reardon, 2007; Reardon & Galindo, 2006) (see Table 3).

A number of other studies confirm that most of the growth in the black-white achievement gap occurs prior to fifth grade. Data from SECCYD suggest the gap is large at the start of kindergarten, and grows in the early elementary grades (Murnane, et al., 2006). NAEP-LTT data show that the black-white math gap (though not the reading gap) widens slightly from age nine to 13 (Ferguson, 1998; Neal, 2005; Phillips, Crouse, & Ralph, 1998). The development of NAEP reading and math gaps from age 13 to 17 is less clear—there is no consistent pattern in the size of the gaps. The gaps generally do not appear to widen much in this period (Ferguson, 1998), but these results are less certain because differential dropout patterns may bias the estimates of the gaps at age 17. NELS data likewise suggest that the gaps change relatively little following eighth grade (LoGerfo, et al., 2006; Neal, 2006). Finally, state-level data from Texas and North Carolina are consistent with this finding; the gap in state test scores in those states grows relatively little in standard deviation units over the latter half of elementary school and middle school (Clotfelter, et al., 2006; Hanushek & Rivkin, 2006).

The Development of Hispanic-White Achievement Gaps

The most detailed evidence on the development of Hispanic-white gaps comes from the ECLS-K, which includes a large sample (roughly 4,000 students) of Hispanic students that can be disaggregated by national origin, generational status, and English proficiency (Reardon & Galindo, 2006). In addition, because the ECLS-K study administered the math test orally in either English or

Spanish, depending on students' language proficiency, ECLS-K estimates of the Hispanic-white math gap are not biased by the changing English proficiency of the students.⁸

ECLS-K data (Table 3) indicate that the Hispanic-white math and reading gaps at the start of kindergarten are very similar in magnitude to the black-white gap; yet math and reading gaps decrease for Hispanic students during elementary school, while the black-white gaps widen during the same period (Fryer & Levitt, 2004, 2005; Reardon & Galindo, 2006). In math, the gap shrinks by 40%, from three-quarters to less than one-half of a standard deviation. Most of the narrowing of the gap occurs in kindergarten and first grade.

Evidence from other studies suggests the Hispanic-white achievement gaps continue to narrow, albeit slowly, through middle and high school. Data from North Carolina indicate that Hispanic students gain ground on white students in both math and reading from third through eighth grade, closing both gaps by over 0.10 standard deviations (Clotfelter, et al., 2006).

Hispanic-white gaps do not develop similarly for all subgroups of the Hispanic student population. Using ECLS-K data, Reardon and Galindo (2006) provide a detailed description of the development of achievement gaps from kindergarten through fifth grade, by Hispanic subgroups (as defined by national origin and generational status). They find that Hispanic-white math and reading gaps appear to narrow during elementary school for all Hispanic subgroups, though the rate and pattern of this narrowing differs across subgroups. Reardon and Galindo speculate that because the rapid gains for Hispanics occur generally in the first two years of schooling, and because they are concentrated among recent immigrants and students with the lowest levels of

⁸ Because the ECLS-K reading test was only administered in English, students not proficient in oral English were not assessed in reading; as a result, estimates of the Hispanic-white reading gap are typically generalizable only to the subsample of Hispanics proficient in oral English at the start of kindergarten (see, for example, Fryer & Levitt, 2004, 2005; Reardon & Galindo, 2006). In math, Hispanic students not proficient in English took the test in Spanish. Although there is evidence that the scores of some Hispanic students are biased downward because they took the math test in English rather than Spanish, this bias likely affects only Hispanic students whose English proficiency was just above the threshold level required for them to take the test in English; it is unlikely to substantially affect the overall pattern of Hispanic math scores (Robinson, 2010).

English proficiency at the start of kindergarten, it is likely that much of the narrowing of the gap in kindergarten and first grade is attributable to the development of English language skills among these students. Because most other studies do not have sufficiently large samples of Hispanic subgroups (or do not collect data on national origin and generational status), we know relatively little about the development of achievement gaps by Hispanic subgroups in middle and high school.

The Development of Asian-White Achievement Gaps

There is relatively little detailed evidence regarding the development of Asian-white achievement gaps. Reardon and Galindo (2006) and Fryer and Levitt (2005) report the development of Asian-white gaps using ECLS-K data for the three-quarters of Asian-origin students proficient in oral English at the start of kindergarten.⁹ These estimates likely overstate the average achievement of Asian-origin students, since the excluded students are generally from recently-immigrated and lower socioeconomic status families. That said, the ECLS-K data (Table 3) show that Asian-origin students who are proficient in oral English at the start of kindergarten appear to have similar levels of math skills as white students through third grade, but better math skills in fifth and eighth grades. In reading, the Asian-origin students in the sample have scores roughly two- to three-tenths of a standard deviation higher than white students through first grade. This advantage narrows to non-significance by fifth grade, but slightly increases to almost two-tenths of a standard deviation in eighth grade (Fryer & Levitt, 2005; Reardon & Galindo, 2006). Cheadle (2008) finds a similar reading advantage for Asian students through first grade, but the advantage narrows rapidly between second and third grade. He finds that the Asian-white gap does not grow during kindergarten, decreases during first grade, and widens in second and third grades. Table 3 includes our estimates of the Asian-white gaps computed from ECLS-K.

⁹ The ECLS-K math test was only administered in English or Spanish (and the reading test was only administered in English), so the roughly 22% of Asian-origin students in the sample who were not proficient in oral English did not take the tests in the Fall of kindergarten.

Given the heterogeneity of the Asian population—in terms of national origin, recency of immigration, socioeconomic status, and context of immigration—generalizations about Asian students as a monolithic group are as problematic as they are for Hispanic students. Certainly more detailed research on Asian-origin students' achievement patterns would be useful.

The Development of Socioeconomic Achievement Gaps

Although considerable research documents the strong association between family socioeconomic characteristics and children's cognitive development and school achievement, there is relatively little such research that provides a descriptive analysis of the development of the socioeconomic achievement gradient over time. Reardon (2011) finds that the income achievement gap changes little as children progress through school. One factor confounding such analyses is the fluid nature of socioeconomic status—unlike race or gender, socioeconomic characteristics of a family change over time, often quite dramatically.

In addition to the findings by Reardon (2011) on the patterns of cognitive development over time, we present here some new analyses using the ECLS-K and NAEP data and add gaps by parental education. Using an indicator of the highest level of education completed by either parent, we begin by examining trends in NAEP-LTT data. Averaging across all years for which we have NAEP data (see Table 1), we see that the average reading achievement gap between students from families with a high school diploma and those with a four-year college degree widens from 0.48 standard deviations at age 13 to 0.55 standard deviations at age 17. The average math gap widens from 0.59 standard deviations to 0.67 standard deviations from age 13 to 17. Of course, if low-achieving students from less educated families leave school before age 17 at higher rates than similarly low-achieving students from families with college-educated parents, then these numbers may underestimate the rate of increase in the socioeconomic gaps as children progress through school.

Data from longitudinal studies such as ECLS-K and NELS also show that socioeconomic gaps widen somewhat as students progress through school, except during the first few years of schooling, when they appear to narrow modestly (Reardon, 2011). Our analysis of ECLS-K data suggests that both math and reading achievement gaps narrow by roughly 10% during the first two years of schooling, but then widen slowly through eighth grade (Table 3).

Do Achievement Gaps Grow Differently Across the Achievement Range?

Most studies examining achievement disparities between groups focus on differences in mean achievement. There are, however, important reasons to examine the disparities across the full distribution of test scores. For example, underlying the debate regarding affirmative action in admissions to highly competitive colleges is the fact that black and Hispanic students are dramatically underrepresented in the upper end of the achievement distribution. Neal (2006, see Figures 2a-2d) shows that roughly 5 percent of black students aged 13-17 years old in the 1990s had math scores in the top quartile of the white math score distribution. This means that black students are underrepresented by 80 percent in the top quartile of the distribution, a finding that has enormous implications for black students' access to elite colleges and employment in jobs with the highest skill demands (and the highest pay). Such patterns suggest the importance of investigating not only differences in the black and white test score distributions, but also of investigating when and how such differences emerge.

Answering this question turns out to be more complex than it would seem, however, because any comparison of the magnitude of gaps or differences in growth rates relies on the assumption that the test metric used is interval-scaled. Clotfelter, Ladd, and Vigdor (2006) investigate whether the gap in scores between the 90th percentiles of the black and white test score distributions grows or narrows faster than the gap between the 10th percentiles of the distributions.

They find that in math, racial test score gaps measured in standard deviation units generally narrow from grades three to eight at the 10th percentiles of the score distributions, and widen at the same time at the 90th percentiles of the distributions. They find no such pattern for reading.

Clotfelter, Ladd, and Vigdor interpret the math pattern as potentially a result of accountability pressures, arguing that the compression of the gap at the low end of the test score distribution is a result of policies that push schools to reduce the percentage of students scoring below certain thresholds. Likewise, they view the expansion of the gap at the high end as a result of the diversion of resources away from high-achieving minority students (because such students are in schools with many low-achieving students). While this is a plausible explanation, it is also possible that the results are an artifact of the tests used to measure the gaps. If the third- and eighth-grade tests are not both scored in interval-scaled metrics, and if the eighth-grade test metric is more sensitive to variation at the high end of the distribution than is the third-grade test, then the pattern they find would be observed in the absence of any true difference in the rate of the gap growth.

Likewise, determining whether achievement gaps grow faster or slower between initially high- and low-achieving students relies on the assumption that test scores are interval-scaled. In addition, measurement error in test scores will also tend to bias such estimates, because conditioning growth rates on scores measured with error will systematically bias estimates of differences in growth rates (Hanushek & Rivkin, 2006; Reardon, 2007). Relatively little empirical research has attempted to systematically address the question of whether achievement gaps within a cohort grow or narrow differentially across the range of skill distribution.¹⁰ What research there is has focused exclusively on the black-white gap, and generally has not adequately addressed the complexities described here. As a result, there is little we can say with certainty to this point.

¹⁰ The comparison across cohorts relies much less on the assumption of interval scaling, since it is possible to compare the full test score distributions across cohorts. See for example, Hedges and Nowell (1999) and Ferguson (1998).

How Much of Racial/Ethnic Achievement Gaps Can Be Explained by Socioeconomic Status?

A relatively common question addressed in studies of racial/ethnic achievement gaps (particularly the black-white gap) is the extent to which the observed gaps can be explained by socioeconomic differences between the groups. Using ECLS-K data, Fryer and Levitt (2005) show that socioeconomic factors explain almost all (85 percent) of the black-white math gap, and all of the reading gap at the start of kindergarten (in fact, they find that the black students score higher in reading than white students of the same socioeconomic status). By third grade, however, they find that the same socioeconomic factors account for only about 60 percent of both the math and reading black-white gaps. This finding suggests that socioeconomic factors explain, in large part, the black-white differences in cognitive skills at the start of formal schooling, but do not account for the growth of the black-white gap as children progress through elementary school. This observation has significant implications for understanding the role of schooling in producing or exacerbating achievement gaps.

The Fryer and Levitt analysis is notable for being the only one that shows that the black-white gap at kindergarten entry can be almost completely accounted for by socioeconomic differences between black and white students. Earlier studies typically found that socioeconomic factors explain roughly half of the black-white gap at kindergarten entry, though these results were based on samples of children that are disproportionately poor (Brooks-Gunn, Klebanov, & Duncan, 1996; Phillips, Brooks-Gunn, et al., 1998). Other recent studies, however, show that socioeconomic factors typically account for even less—between 25 and 40 percent—of the black-white math and reading gaps through middle and high school.¹¹ Using the same ECLS-K data as Fryer and Levitt but

¹¹ One possible reason why Fryer and Levitt's model explains more of the black-white gap in elementary school than other studies may be that they use a somewhat larger list of covariates than other studies, including age, birthweight, gender, number of children's books (and its quadratic term), mother's age at first birth, WIC receipt, and a composite socioeconomic status variable that includes family income, parental education, and parental occupation measures (Fryer & Levitt, 2006).

a different version of the test score metric (the unstandardized scale scores), Murnane and colleagues (2006) report that the socioeconomic differences between white and black students accounts for only one-third of the math gap and 15 percent of the reading gap by third grade (though they find that socioeconomic factors fully account for the black-white gap in kindergarten).¹² Similarly, using the SECCYD data, they find that socioeconomic status accounts for one-third of the math and one-quarter of the reading black-white gaps in third grade (Murnane, et al., 2006). Phillips, Crouse, & Ralph's (1998) analysis of Prospects and NELS data shows SES explains about 35 to 40 percent of black-white achievement gaps in math, reading, and vocabulary across the span of second- through twelfth-grade. North Carolina data show similar patterns, with socioeconomic factors (as well as region of the state and urbanicity) accounting for about 35 percent of the math and reading gaps from third through eighth grades (Clotfelter, et al., 2006).

Yet, recent research suggests that socioeconomic factors may explain an even larger share of the black-white achievement gap than estimated by Fryer and Levitt. Rothstein and Wozny (2013) argue that parent-reported current annual income (which is what Fryer and Levitt and most studies use) is a noisy measure of families' long-run income; the use of a noisy measure of long-run income leads to underestimation of the extent to which race/ethnicity achievement gaps are explained by socioeconomic status. When Rothstein and Wozny take long-run income into account, they estimate that the black-white math gap in 3rd grade is 0.15 standard deviations, considerably smaller than Fryer and Levitt's (2006) estimate of 0.38 standard deviations.

Other research using data from the National Longitudinal Survey of Youth suggests that the black-white achievement gap in adolescence can be entirely explained by models that account for multigenerational measures of socioeconomic status (Mandara, Varner, Greene, & Richman, 2009).

¹² Murnane et al.'s results differ substantially from those of Fryer and Levitt likely because Murnane et al. report the predicted black-white gap from a random-coefficient growth model while Fryer and Levitt report results from a series of repeated cross-sectional covariate adjustment models (Fryer & Levitt, 2006; Murnane, et al., 2006). We prefer the Fryer and Levitt results because they rely less on assumptions regarding interval scaling of the test metric and the stable effects of socioeconomic factors over time.

Mandara and colleagues found that, conditional on parental and grandparental socioeconomic status and parenting practices, black students actually have slightly higher test scores than white students. Thus, research accounting for long-run parent income (e.g., Rothstein & Wozny, 2013) or multigenerational socioeconomic status (e.g., Mandara, et al., 2009) tends to explain a much greater proportion of black-white achievement gaps than does research using current, single-generational measures.

The evidence regarding the extent to which socioeconomic factors account for Hispanic-white achievement gaps is less mixed: Research consistently finds that these gaps are largely or entirely explained by socioeconomic status. For example, using ECLS-K data, Fryer and Levitt show that socioeconomic factors account for 75 to 85 percent of the Hispanic-white gaps in kindergarten and 85 to 100 percent of the gaps in third grade (Fryer & Levitt, 2006). One explanation for this different pattern may be that the Hispanic-white gaps in kindergarten are partly due to Hispanic students' lower levels of English proficiency; as students progress through school, their English skills improve, the gaps narrow (see Table 3), so that socioeconomic factors explain an increasing proportion of the gaps. Similarly, Reardon and Galindo (2006; tables B11, B12) show that Hispanic-white gaps, conditional on socioeconomic status, narrow from kindergarten through fifth grade, while the corresponding black-white gaps widen at the same time. North Carolina data show a similar pattern: Hispanic-white socioeconomic-adjusted math and reading gaps are small (one-tenth of a standard deviation) in third grade, non-existent by fifth grade, and reverse sign by eighth grade, when Hispanic students score higher, on average, in both math and reading than socioeconomically similar white students (Clotfelter, et al., 2006).

Do Achievement Gaps Grow Between or Within Schools?

A central question in understanding test score gaps is the extent to which such gaps can be attributed to differences in average school quality between schools attended by students of different racial, ethnic, or socioeconomic groups. If, for example, black students attend, on average, lower quality schools than white students, we would expect the between-school component of the black-white achievement gap to grow over time. If black and white students receive unequal instructional opportunities when attending the same schools, we would expect the within-school component of the black-white gap to grow over time. Of course, it is difficult to disentangle the effects of school quality from the sorting processes that produce racially and socio-economically segregated schools and that may result in lower-ability students, regardless of race or socioeconomic status, attending schools that have lower proportions of white or middle-class students. Likewise, it is not clear that differences in achievement gains can be attributed solely to schooling processes (particularly given the evidence that the gaps predate kindergarten; see Fryer & Levitt, 2004; Lee & Burkham, 2002; Reardon & Galindo, 2006), given unequal family resources, neighborhood context, and opportunity structures (which may lead to unequal motivation even in the presence of equal home and school resources). Nonetheless, an understanding of the relative contribution of between- and within-school factors, as well as of family background and out-of-school social context, is essential for determining the appropriate policy remedies for the gaps.

Despite the importance of disentangling between- and within-school patterns in the growth of achievement gaps, there is relatively little clear evidence on this point. Fryer and Levitt (2005) find that the black-white gap is small in kindergarten, net of family socioeconomic characteristics, but grows from kindergarten through third grade (using ECLS-K data)—a pattern that suggests that observable family background characteristics are not solely responsible for the post-kindergarten growth in the gap. However, they also find that the black-white gap grows through third grade even between black and white students attending the same school, a finding they interpret to mean that

between-school differences in school quality do not account for much of the growth of the black-white gap during elementary school (but see Hanushek & Rivkin, 2006; see also Reardon, 2007, for an explanation of discrepancies between Fryer & Levitt's and Hanushek & Rivkin's analyses).

Another approach to studying whether gaps develop within or between schools is to investigate the extent to which beneficial instructional practices and teacher attributes are differentially distributed across different groups of students within and between schools. For example, using national data (ECLS-K), Desimone and Long (2010) found that less experienced teachers were more likely to be assigned to minority students and students from low-income families. These patterns are certainly partly due to differences among schools in average teacher experience levels, differences that are associated with school racial composition. However, more recent research provides evidence that teacher sorting occurs within schools as well as between schools. Kalogrides and colleagues found that novice teachers are assigned to lower-achieving, minority, and poor students within the same school, potentially exacerbating achievement gaps (Kalogrides & Loeb, forthcoming; Kalogrides, Loeb, & Beteille, 2013). Another possible source of widening achievement gaps is growing racial segregation (Berends & Peñaloza, 2010; Condrón, Tope, Steidl, & Freeman, 2013). In sum, the literature suggests that a number of factors—both within and across schools—likely contribute to racial achievement gaps.

What Are the Labor Market Effects of Achievement Gaps?

From a labor market perspective, achievement disparities are important primarily because test score disparities in elementary and secondary school are highly predictive of corresponding disparities in subsequent labor market outcomes. Data from the Current Population Survey (CPS) from 2011 and 2012 show that the median black and Hispanic male full-time workers earn 76% and 67% of what the median white full-time male worker earns. For females, the median black and

Hispanic full-time workers earn 84% and 73% of their white counterparts. These wage differentials have remained stable for the past decade.

A sizeable body of research has investigated the extent to which these wage disparities are attributable to differences in cognitive skill obtained prior to entering the labor force (i.e., in childhood, elementary school, and secondary school), typically measured by cognitive achievement test scores. In general, this research finds that roughly one-half to all of the male black-white wage gap can be accounted for by black-white differences in human capital, as proxied by scores on the Armed Forces Qualification Test (AFQT) when individuals were near completing high school (Bollinger, 2003; Carneiro, Heckman, & Masterov, 2003; Neal & Johnson, 1996).¹³ Although there is much less evidence regarding the extent to which test score differences account for Hispanic-white wage gaps, evidence from NLSY suggests that test score differences account for virtually all of the male Hispanic-white wage gap (Carneiro, et al., 2003).¹⁴ These findings suggest that factors such as racial discrimination in the labor market are not responsible for much, if any, of the black-white wage differential, a conclusion supported by other research that compares earnings differentials among different cohorts. For example, O'Neill (1990) reviews possible explanations for the decrease in the black-white wage differential during the twentieth century, and finds that differences in the educational and workplace experiences of black and white men are much more important in explaining the wage gap between the races than is labor market discrimination. Examining hourly wages six years after high school graduation, Murnane, Willett, and Levy (1995) find that cognitive skills predict labor market outcomes more in the mid-1980s than they did in the late-1970s, suggesting the increasing importance of cognitive skills in wage determination.

¹³ With regard to wage gaps for women, the evidence is less clear because of differential selection into the labor force among women. Among women in the labor force, however, black and Hispanic women earn, on average, the same or more than white women after adjusting for AFQT scores (Bollinger, 2003; Carneiro, et al., 2003).

¹⁴ Note that the NLSY Hispanic sample is made up of Hispanics living in the U.S. in 1979, so this finding may not be representative of the experiences of the current U.S. Hispanic population.

It is worth noting, however, that not all gains in reducing black-white wage inequality during the twentieth century can be attributed to narrowing skill gaps (O'Neill, 1990). For instance, the relative earnings of blacks increased during the 1960s and 1970s—arguably due to governmental intervention in reducing employment discrimination—for cohorts who had already completed schooling (Donohue & Heckman, 1991). Likewise, some research suggests that labor market discrimination remains an important factor in labor market outcomes (see, for example, Grodsky & Pager, 2001; Pager, 2003), though even this research acknowledges that the majority of the black-white wage differential can be attributed to black-white differences in human capital obtained prior to entering the labor market.

The finding that human capital differences account for a large portion of black-white wage gaps suggests that efforts to reduce black-white achievement disparities by the end of adolescence may substantially reduce subsequent black-white wage gaps, though such a reduction would take decades to take full effect. Even an immediate elimination of the gaps in human capital among adolescents would not, presumably, reduce wage disparities among cohorts already in the labor force. As a result, wage differentials among older cohorts may persist for decades, until these cohorts age out of the labor force.¹⁵

In addition to concerns regarding the magnitude of the differences in mean test scores among individuals of different racial groups, a number of researchers have called attention to the effects of racial disparities at the upper end of the achievement distribution (Hanushek & Rivkin,

¹⁵ In addition, we might expect that—to the extent that any of the current wage differentials is due to employers' "statistical discrimination" practices (i.e., employers' use of race or education levels as a proxy for unobserved skill levels believed to differ systematically between racial groups)—an immediate reduction in achievement gaps may not yield immediate wage gap reductions even among the youngest cohorts, because it may take time for employers to adjust their prior beliefs about racial differences in human capital. However, Altonji and Pierret (2001) find little evidence for statistical discrimination on the basis of race, net of educational attainment. Moreover, they find also that although employers may statistically discriminate on the basis of employees' educational attainment in assigning initial wage levels, employees' productivity plays a much larger role in wage determination as the employee accrues more experience—employers appear to adjust their prior beliefs, at least with regard to their own employees, as they gain more information. These findings suggest that the impact of a reduction in the black-white skill gap may have a relatively immediate effect on wage gaps among cohorts for whom the skill gap is reduced.

2007; Neal, 2006). Moreover, some evidence indicates that the increase in the returns to education in the 1980s was largest for those in the top quartile of the achievement distribution (Heckman & Vytlačil, 2001). Because whites are substantially overrepresented in the highest quartile of the achievement distribution, this pattern suggests that racial disparities at the top of the achievement distribution have become increasingly salient in shaping labor market inequality.

Finally, although much of the economics literature focuses on the role of cognitive skills in explaining wage gaps, there is an emerging emphasis on the role of so-called “non-cognitive” skills (e.g., motivation, self-control, perseverance) on wages. For example, a recent study using data from the Perry Preschool experiment examined the extent to which both cognitive and non-cognitive benefits from the experimental condition (focused on children from low-income families) affected income at age 27 (Heckman, Pinto, & Savelyev, 2012). The analyses revealed that, for males, both cognitive and non-cognitive factors significantly predicted differences between the treatment and control groups in monthly income at age 27 and whether the individual was employed at age 40. In addition, treatment effects in non-cognitive factors associated with reduced externalizing problem behaviors at ages 7 to 9 (e.g., lying, disrupting class, acting aggressive to peers) were associated with treatment differences in other adult outcomes such as drug-use and arrests, factors that are likely undesirable to employers. This evidence suggests that non-cognitive skills have an effect on later income, independent of cognitive skills. Moreover, other evidence suggests that non-cognitive skills are associated with achievement (Duckworth & Seligman, 2005) and socioeconomic status (Moffitt et al., 2011; Noble, McCandliss, & Farah, 2007), suggesting that non-cognitive skill gaps in childhood may contribute to academic achievement gaps as well. Thus, interventions might consider a focus on reducing non-cognitive gaps to partially address both cognitive gaps and later wage gaps (Lindqvist & Vestman, 2011).

Conclusion

In this review, we have attempted to summarize the state of knowledge regarding racial/ethnic achievement gaps and to suggest areas in which more research is needed. Despite the complexity of answering questions about achievement gaps due to the need to rely on imperfect measures of cognitive skills, several key patterns are evident from our summary of the research.

First, racial/ethnic achievement gaps are narrower now than they were 30-40 years ago; this is particularly true for the black-white achievement gaps. Second, these gaps remain quite large today, ranging from 0.5 to 1.0 standard deviations. Third, the black-white gap appears to widen during the school years—particularly in early elementary school—in ways that are not explained by socioeconomic family background characteristics, a pattern that suggests that schooling appears to contribute to the growth of the gaps. The same patterns are not found for Hispanic-white and Asian-white disparities, however; for these groups, socioeconomic differences account for a large portion of the gaps, and processes of second-language acquisition appear to contribute to a narrowing of the gap as children progress through school. Fourth, achievement disparities have large effects in the labor market, and explain a large portion of racial/ethnic income disparities.

Less clear in extant research are the processes and mechanisms that produce racial and ethnic achievement disparities. Although it is clear that family background and schooling each play some role in the development of achievement gaps, we do not have good evidence on exactly how—or how much—each contributes. Nor do we have any good evidence regarding the extent to which these processes may vary among racial or ethnic groups. Most importantly, without a good understanding of the mechanisms that produce these gaps, we have very little evidence regarding how we might reduce them. In addition to documenting the magnitude, trends, and development of achievement gaps, we need more and better research regarding the effectiveness of social and educational policy to reduce them.

References

- Altonji, J. G., & Pierret, C. R. (2001). Employer learning and statistical discrimination. *Quarterly Journal of Economics*, 116(1), 313-350.
- Benson, J., & Borman, G. D. (2010). Family, neighborhood, and school settings across seasons: When do socioeconomic context and racial composition matter for the reading achievement growth of young children? *Teachers College Record*, 112(5), 1338-1390.
- Berends, M., & Peñaloza, R. V. (2010). Increasing Racial Isolation and Test Score Gaps in Mathematics: A 30-Year Perspective *Teachers College Record*, 112(4), 978-1007.
- Bollinger, C. (2003). Measurement error in human capital and the black-white wage gap. *The Review of Economics and Statistics*, 85(3), 578-585.
- Bond, T. N., & Lang, K. (2013). The Black-White Education-Scaled Test-Score Gap In Grades K-7. *National Bureau of Economic Research Working Paper*, 19243.
- Brooks-Gunn, J., Klebanov, P. K., & Duncan, G. J. (1996). Ethnic differences in children's intelligence test scores: Role of economic deprivation, home environment, and maternal characteristics. *Child Development*, 67(2), 396-408.
- Card, D., & Rothstein, J. (2006). Racial segregation and the black-white test score gap. *National Bureau of Economic Research Working Papers*, 12078.
- Carneiro, P., Heckman, J. J., & Masterov, D. V. (2003). *Labor market discrimination and racial differences in premarket factors* (No. w10068). Cambridge, MA: National Bureau of Economic Research.
- Cheadle, J. E. (2008). "Educational Investment, Family Context, and Children's Math and Reading Growth from Kindergarten through Third Grade. *Sociology of Education*, 81(1), 1-31.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). *The academic achievement gap in grades three to eight* (Working Paper No. 12207). Cambridge, MA: National Bureau of Economic Research.
- Condrón, D., Tope, D., Steidl, C. R., & Freeman, K. J. (2013). Racial segregation and the black/white achievement gap 1992-2009. *The Sociological Quarterly*, 54(1), 130-157.
- Desimone, L. M., & Long, D. (2010). Teacher Effects and the Achievement Gap: Do Teacher and Teaching Quality Influence the Achievement Gap Between Black and White and High- and Low-SES Students in the Early Grades? *Teachers College Record*, 112(12), 324-373.
- Donohue, J., & Heckman, J. J. (1991). Continuous versus episodic change: The impact of civil rights policy on the economic status of blacks. *Journal of Economic Literature*, 29, 1603.
- Duckworth, A. L., & Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological science*, 16(12), 939-944.
- Ferguson, R. F. (1998). Test-Score Trends Along Racial Lines, 1971 to 1996: Popular Culture and Community Academic Standards. In N. J. Smelser, W. J. Wilson & F. Mitchell (Eds.), *America Becoming: Racial Trends and Their Consequences* (Vol. 1, pp. 348-390). Washington, D.C.: National Academies Press.
- Fryer, R. G., & Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *The Review of Economics and Statistics*, 86(2), 447-464.
- Fryer, R. G., & Levitt, S. D. (2005). *The black-white test score gap through third grade* (Working paper No. w11049). Cambridge, MA: National Bureau of Economic Research.
- Fryer, R. G., & Levitt, S. D. (2006). The Black-White Test Score Gap through Third Grade. *American Law and Economics Review*, 8(2), 249-281.
- Grissmer, D. W., Flanagan, A., & Williamson, S. (1998). Why did the Black-White score gap narrow in the 1970s and 1980s? In C. Jencks & M. Phillips (Eds.), *The Black-White Test Score Gap* (pp. 182-228). Washington, D.C.: Brookings Institution Press.

- Grodsky, E., & Pager, D. (2001). The structure of disadvantage: Individual and occupational determinants of the black-white wage gap. *American Sociological Review*, 66(August), 542-567.
- Guryan, J. (2004). Desegregation and Black dropout rates. *American Economic Review*, 94(4), 919-943.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage Publications.
- Hanushek, E. A., & Rivkin, S. G. (2006). *School quality and the black-white achievement gap*. Unpublished manuscript.
- Hanushek, E. A., & Rivkin, S. G. (2007). *School quality and the black-white achievement gap*. Unpublished manuscript.
- Heckman, J. J., Pinto, R., & Savelyev, P. A. (2012). *Understanding the mechanisms through which an influential early childhood program boosted adult outcomes*: National Bureau of Economic Research.
- Heckman, J. J., & Vytlacil, E. (2001). Identifying the role of cognitive ability in explaining the level of and change in the return to schooling. *Review of Economics and Statistics*, 83(1), 1-12.
- Hedges, L. V., & Nowell, A. (1998). Black-White Test Score Convergence Since 1965. In C. Jencks & M. Phillips (Eds.), *The Black-White Test Score Gap* (pp. 149-181). Washington, D.C.: Brookings Institution Press.
- Hedges, L. V., & Nowell, A. (1999). Changes in the black-white gap in achievement test scores. *Sociology of Education*, 72(2), 111-135.
- Ho, A. D., & Haertel, E. H. (2006). *Metric-Free Measures of Test Score Trends and Gaps with Policy-Relevant Examples* (CSE Report No. 665). Los Angeles, CA: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies.
- Ho, A. D., & Reardon, S. F. (2012). Estimating Achievement Gaps From Test Scores Reported in Ordinal "Proficiency" Categories. *Journal of Educational and Behavioral Statistics*, 37(4), 489-517.
- Johnson, R. C. (2011). Long-Run Impacts of School Desegregation and School Quality on Adult Attainments. *National Bureau of Economic Research Working Papers*, 16664.
- Kalogrides, D., & Loeb, S. (forthcoming). Different teachers, different peers: The magnitude of student sorting within schools. *Educational Researcher*.
- Kalogrides, D., Loeb, S., & Beteille, T. (2013). Systematic Sorting Teacher Characteristics and Class Assignments. *Sociology of Education*, 86(2), 103-123.
- Lee, V. E., & Burkham, D. T. (2002). *Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School*. Washington, DC: Economic Policy Institute.
- Lindqvist, E., & Vestman, R. (2011). The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment. *American Economic Journal: Applied Economics*, 3(1), 101-128.
- LoGerfo, L., Nichols, A., & Reardon, S. F. (2006). *Achievement Gains in Elementary and High School*. Washington, DC: Urban Institute.
- Lord, F. M. (1980). *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lord, F. M., & Novick, M. R. (1968). *Statistical Theory of Mental Test Scores*. Reading, MA: Addison Wesley.
- Mandara, J., Varner, F. A., Greene, N., & Richman, S. B. (2009). Family predictors of the black-white achievement gap in adolescence. *Journal of Educational Psychology*, 101(4), 867-878.
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, 108(7), 2693-2698.

- Murnane, R. J., Willett, J. B., Bub, K. L., & McCartney, K. (2006). Understanding trends in the black-white achievement gaps during the first years of school. *Brookings-Wharton Papers on Urban Affairs*.
- Murnane, R. J., Willett, J. B., & Levy, F. (1995). The Growing Importance of Cognitive Skills in Wage Determination. *Review of Economics and Statistics*, 78(2), 251-266.
- National Council on Measurement in Education, & Brennan, R. L. (Eds.). (2006). *Educational Measurement* (4th ed.): Praeger.
- Neal, D. A. (2005). Why has Black-White skill convergence stopped? Unpublished manuscript. University of Chicago.
- Neal, D. A. (2006). Why has Black-White skill convergence stopped? In E. A. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education* (Vol. 1, pp. 511-576). New York: Elsevier.
- Neal, D. A., & Johnson, W. R. (1996). The role of premarket factors in black-white wage differences. *The Journal of Political Economy*, 104(5), 869-895.
- Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental science*, 10(4), 464-480.
- O'Neill, J. (1990). The role of human capital in earnings differentials between black and white men. *Journal of Economic Perspectives*, 4(4), 25-45.
- Pager, D. (2003). The mark of a criminal record. *American Journal of Sociology*, 108(5), 937-975.
- Phillips, M., Brooks-Gunn, J., Duncan, G. J., Klebanov, P., & Crane, J. (1998). Family Background, Parenting Practices, and the Black-White Test Score Gap. In C. Jencks & M. Phillips (Eds.), *The Black-White Test Score Gap* (pp. 103-148). Washington, D.C.: Brookings Institution Press.
- Phillips, M., Crouse, J., & Ralph, J. (1998). Does the black-white test score gap widen after children enter school? In C. Jencks & M. Phillips (Eds.), *The black-white test score gap* (pp. 229-272). Washington, DC: Brookings Institution Press.
- Reardon, S. F. (2007). *Thirteen ways of looking at the black-white test score gap*. Unpublished manuscript.
- Reardon, S. F. (2011). The widening socioeconomic status achievement gap: New evidence and possible explanations. In R. J. Murnane & G. J. Duncan (Eds.), *Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children*. New York: Russell Sage Foundation.
- Reardon, S. F., & Galindo, C. (2006). *Patterns of Hispanic Students' Math and English Literacy Test Scores in the Early Elementary Grades*: National Task Force on Early Childhood Education for Hispanics.
- Reardon, S. F., Valentino, R., Kalogrides, D., Shores, K. A., & Greenberg, E. H. (2013). Patterns and Trends in Racial Academic Achievement Gaps Among States, 1999-2011. Unpublished Working Paper. Center for Education Policy Analysis, Stanford University.
- Robinson, J. P. (2010). The Effects of Test Translation on Young English Learners' Mathematics Performance. *Educational Researcher*, 39(8), 582-590.
- Robinson, J. P., & Lubienski, S. T. (2011). The Development of Gender Achievement Gaps in Mathematics and Reading During Elementary and Middle School Examining Direct Cognitive Assessments and Teacher Ratings. *American Educational Research Journal*, 48(2), 268-302.
- Rothstein, J., & Wozny, N. (2013). Permanent Income and the Black-White Test Score Gap. *The Journal of Human Resources*, 48(3), 509-544.
- Selzer, M. H., Frank, K. A., & Bryk, A. S. (1994). The metric matters: the sensitivity of conclusions about growth in student achievement to choice of metric. *Educational Evaluation and Policy Analysis*, 16(1), 41-49.
- Vigdor, J., & Ludwig, J. (2007). Segregation and the black-white test score gap. *National Bureau of Economic Research Working Paper*, 12988.

Yen, W. M., & Fitzpatrick, A. R. (2006). Item Response Theory. In National Council on Measurement in Education & R. L. Brennan (Eds.), *Educational Measurement* (4th ed., pp. 111-153): Praeger.

Table 1: Estimated Achievement Gaps, NAEP Long-Term Trend Assessments, by Group, Subject, Age, and Test Year

Test Year	Black-White Gap						Hispanic-White Gap						HS-College Degree Gap			
	Reading			Math			Reading			Math			Reading		Math	
	9-y.o.	13-y.o.	17-y.o.	9-y.o.	13-y.o.	17-y.o.	9-y.o.	13-y.o.	17-y.o.	9-y.o.	13-y.o.	17-y.o.	13-y.o.	17-y.o.	13-y.o.	17-y.o.
1971	-1.04	-1.08	-1.15													
1975	-0.92	-1.01	-1.19				-0.88	-0.83	-0.92							
1978				-0.88	-1.08	-1.08				-0.59	-0.86	-0.85			-0.53	-0.66
1980	-0.84	-0.91	-1.19				-0.82	-0.78	-0.75				-0.55	-0.57		
1982				-0.83	-1.02	-0.98				-0.57	-0.66	-0.83			-0.58	-0.59
1984	-0.78	-0.74	-0.79				-0.75	-0.64	-0.68				-0.42	-0.53		
1986				-0.75	-0.79	-0.93				-0.63	-0.63	-0.79			-0.56	-0.67
1988	-0.71	-0.53	-0.55				-0.58	-0.61	-0.64				-0.36	-0.48		
1990	-0.79	-0.58	-0.71	-0.81	-0.88	-0.67	-0.62	-0.68	-0.53	-0.65	-0.70	-0.84	-0.43	-0.47	-0.57	-0.72
1992	-0.83	-0.73	-0.86	-0.82	-0.93	-0.86	-0.65	-0.69	-0.61	-0.70	-0.63	-0.65	-0.48	-0.49	-0.63	-0.61
1994	-0.80	-0.77	-0.67	-0.74	-0.90	-0.89	-0.79	-0.75	-0.73	-0.81	-0.77	-0.71	-0.44	-0.55	-0.59	-0.74
1996	-0.74	-0.82	-0.69	-0.75	-0.92	-0.89	-0.64	-0.71	-0.70	-0.66	-0.81	-0.71	-0.46	-0.60	-0.51	-0.64
1999	-0.90	-0.74	-0.73	-0.82	-0.99	-1.02	-0.71	-0.60	-0.57	-0.76	-0.73	-0.72	-0.48	-0.58	-0.67	-0.57
2004	-0.70	-0.65	-0.61	-0.69	-0.87	-0.88	-0.65	-0.61	-0.49	-0.46	-0.67	-0.63	-0.44	-0.53	-0.54	-0.69
2008	-0.65	-0.56	-0.65	-0.76	-0.81	-0.88	-0.55	-0.69	-0.58	-0.47	-0.65	-0.70	-0.51	-0.55	-0.55	-0.64
2012	-0.60	-0.62	-0.62	-0.71	-0.80	-0.84	-0.57	-0.55	-0.50	-0.48	-0.61	-0.63	-0.69	-0.69	-0.73	-0.84

Source: Authors' calculations from NAEP-LTT data, available at <http://nces.ed.gov/nationsreportcard/ltt/>. Note: Gaps measured in standard deviation units, using the standard deviation across all students for a given test-age-year combination. NAEP-LTT does not report parental education level for 9-year-olds, so we report parental education gaps only for 13- and 17-year-olds.

Table 2: Estimated Achievement Gaps, Main NAEP Assessments, by Group, Subject, Age, and Test Year

Test year	Black-White Gap						Asian-White Gap						HS-College Degree Gap			
	Math			Reading			Math			Reading			Math		Reading	
	4th	8th	12th	4th	8th	12th	4th	8th	12th	4th	8th	12th	8th	12th	8th	12th
1990	-1.02	-0.91	-0.89				0.15	0.15	0.30				-0.54	-0.64		
1992	-1.09	-1.10	-0.89	-0.91	-0.83	-0.74	0.12	0.37	0.21	-0.23	0.04	-0.21	-0.64	-0.67	-0.56	-0.57
1994				-0.94	-0.83	-0.78				-0.10	-0.05	-0.41			-0.49	-0.58
1996	-1.11	-1.09	-0.99				-0.11		-0.11				-0.58	-0.67		
1998				-0.82	-0.75	-0.71				-0.26	-0.18	-0.26			-0.54	-0.55
2000	-0.99	-1.04	-0.94	-0.82			0.12	0.22		0.01			-0.65	-0.70		
2002				-0.83	-0.80	-0.67				-0.12	-0.18	-0.18			-0.52	-0.50
2003	-0.96	-0.98		-0.82	-0.78		0.11	0.09		-0.07	-0.06		-0.58		-0.54	
2005	-0.92	-0.93	-0.91	-0.81	-0.80	-0.68	0.17	0.18	0.15	-0.01	0.00	-0.14	-0.62	-0.69	-0.56	-0.58
2007	-0.91	-0.88		-0.76	-0.78		0.18	0.16		0.04	-0.04		-0.63		-0.57	
2009	-0.90	-0.88	-0.89	-0.73	-0.77	-0.71	0.23	0.22	0.42	0.13	0.04	0.06	-0.67	-0.67	-0.59	-0.58
2011	-0.86	-0.86		-0.71	-0.73		0.24	0.25		0.11	0.05		-0.67		-0.61	
2013	-0.87	-0.84		-0.71	-0.75		0.26	0.34		0.09	0.12		-0.70		-0.67	

Test year	Hispanic-White Gap						Mexican-White Gap					
	Math			Reading			Math			Reading		
	4th	8th	12th	4th	8th	12th	4th	8th	12th	4th	8th	12th
1990	-0.61	-0.66	-0.66				-0.64	-0.76	-0.65			
1992	-0.79	-0.77	-0.56	-0.77	-0.73	-0.58	-0.77	-0.84	-0.59	-0.65	-0.76	-0.61
1994				-0.86	-0.65	-0.63				-0.76	-0.66	-0.68
1996	-0.80	-0.79	-0.73				-0.81	-0.83	-0.84			
1998				-0.83	-0.77	-0.57				-0.79	-0.78	-0.54
2000	-0.85	-0.82	-0.68	-0.84			-0.80	-0.81	-0.73	-0.80		
2002				-0.77	-0.76	-0.53				-0.73	-0.79	-0.59
2003	-0.76	-0.79		-0.76	-0.76		-0.67	-0.77		-0.69	-0.78	
2005	-0.72	-0.74	-0.71	-0.73	-0.71	-0.54	-0.66	-0.71	-0.74	-0.68	-0.72	-0.56
2007	-0.74	-0.73		-0.72	-0.72		-0.69	-0.72		-0.67	-0.73	
2009	-0.72	-0.73	-0.68	-0.71	-0.69	-0.59	-0.67	-0.70	-0.68	-0.68	-0.71	-0.62
2011	-0.68	-0.65		-0.68	-0.63		-0.67	-0.63		-0.69	-0.66	
2013	-0.65	-0.61		-0.67	-0.60		-0.63	-0.62		-0.67	-0.62	

Source: Authors' calculations from Main NAEP data, available at <http://nces.ed.gov/nationsreportcard/naepdata/>. Note: Gaps measured in standard deviation units, using the standard deviation across all students for a given test-grade-year combination. Main NAEP does not report parental education level for 4th-grade students, so we report parental education gaps only for 8th- and 12th-grade students.

Table 3: Estimated Achievement Gaps, Early Childhood Longitudinal Study:1998 Kindergarten Cohort, by Group, Subject, and Grade

	Math						Reading					
	Fall K	Spring K	Spring 1	Spring 3	Spring 5	Spring 8	Fall K	Spring K	Spring 1	Spring 3	Spring 5	Spring 8
<u>Race-based gaps (reference group: White)</u>												
Asian	0.07+	0.05+	0.01+	0.08+	0.29	0.38	0.22	0.32	0.33	0.11+	0.08+	0.17
Black	-0.73	-0.80	-0.77	-0.90	-0.99	-1.00	-0.53	-0.58	-0.54	-0.77	-0.87	-0.95
Hispanic	-0.76	-0.65	-0.53	-0.50	-0.45	-0.44	-0.48	-0.26	-0.26	-0.33	-0.34	-0.36
Other	-0.42	-0.29	-0.34	-0.40	-0.30	-0.33	-0.41	-0.28	-0.30	-0.37	-0.32	-0.33
<u>Parent-education-based gaps (reference group: College degree or higher)</u>												
No HS diploma	-1.40	-1.31	-1.23	-1.25	-1.27	-1.33	-1.26	-1.17	-1.16	-1.39	-1.36	-1.44
HS diploma	-0.90	-0.86	-0.79	-0.83	-0.92	-1.00	-0.88	-0.80	-0.74	-0.84	-0.94	-0.95
Some college	-0.55	-0.50	-0.48	-0.49	-0.53	-0.59	-0.59	-0.46	-0.43	-0.53	-0.55	-0.64
<u>SES-based gaps (reference group: top 20%)</u>												
0-20%	-1.37	-1.31	-1.20	-1.30	-1.33	-1.39	-1.27	-1.17	-1.10	-1.28	-1.35	-1.39
20-40%	-0.84	-0.77	-0.75	-0.79	-0.88	-0.95	-0.90	-0.75	-0.65	-0.76	-0.88	-0.98
40-60%	-0.58	-0.49	-0.46	-0.58	-0.60	-0.65	-0.60	-0.49	-0.41	-0.54	-0.59	-0.68
60-80%	-0.27	-0.27	-0.22	-0.25	-0.28	-0.31	-0.29	-0.25	-0.16	-0.23	-0.28	-0.34

Source: Early Childhood Longitudinal Study—Kindergarten Class of 1998-99 public-use K-8 longitudinal data set (www.nces.ed.gov/ecls). Gap estimates are based on the 4th scaling of test scores (i.e., wave-7 version). Observations were weighted by full longitudinal sample wgt (c1_7cw0). Standard errors are adjusted for strata and PSU. Gaps marked with "+" are not statistically different from zero ($p > .05$); all other estimates in the table are significantly different from zero ($p < .05$).