Sociology of Education
2015, 88(2) 120–139

DOI: 10.1177/0038040715573027
http://soe.sagepub.com

Kindergarten Black–White Test Score Gaps: Re-examining the Roles of Socioeconomic Status and School Quality with New Data

David M. Quinn[1]

Abstract

Black–white test score gaps form in early childhood and widen over elementary school. Sociologists have debated the roles that socioeconomic status (SES) and school quality play in explaining these patterns. In this study, I replicate and extend past research using new nationally representative data from the Early Childhood Longitudinal Study–Kindergarten Class of 2010–2011. I find black–white test score gaps at kindergarten entry in 2010 in reading (*SD* = .32), math (*SD* = .54), and working memory (*SD* = .52 among children with valid scores). Math and reading gaps widened by approximately .06 standard deviations over kindergarten, but the working memory gap was constant. Multivariate regressions show that student SES explained the reading gap at school entry, but gap decompositions suggest that school quality differences were responsible for the widening of the reading gap over kindergarten. SES explained much of the math gap at school entry, but the widening of the math gap could not be explained by SES, school quality, or other hypotheses.

Keywords

black–white test score gap, achievement gap, school quality, educational inequality, socioeconomic status, gap trends

The black–white test score gap is a persistent challenge in our educational landscape. While math and reading gaps have narrowed considerably from their levels in the 1970s (Hedges and Nowell 1999; Reardon, Robinson-Cimpian, and Weathers 2015), large gaps remain. The 2012 National Assessment of Educational Progress–Long-term Trend test (NAEP-LTT) showed gaps of .62 standard deviations in reading and .84 standard deviations in math among 17-year-olds (Reardon et al. 2015). The increasingly larger role that education is playing in social stratification (Fischer and Hout 2006) brings additional urgency to understanding the sources of these gaps.

Since the "Coleman Report" (Coleman et al. 1966), sociologists have debated the extent to which black–white gaps can be explained by students' socioeconomic status (SES) or by differences in the quality of schools attended by black students and white students. The Coleman Report

[1]Harvard Graduate School of Education, Cambridge, MA, USA

Corresponding Author:
David M. Quinn, Harvard Graduate School of Education, 13 Appian Way, Cambridge, MA 02138, USA.
Email: david_quinn@mail.harvard.edu

led some to conclude that families, not schools, make the difference in students' academic achievement (Murnane et al. 2006). Data from the 1980s showing large black–white test score gaps in early childhood, before students had been exposed to much formal schooling, also suggested a substantial role of nonschool factors in gap formation. In these early multivariate analyses, large portions of the gaps went unexplained. More recent work by Fryer and Levitt (2004, 2006) is noteworthy for showing—using data from the Early Childhood Longitudinal Study–Kindergarten Class of 1998–1999 (ECLS-K:1999)—that black–white gaps at kindergarten entry could be either entirely or almost entirely explained by SES. However, these analyses also showed that gaps widened as students progressed through elementary school, and this widening could not be explained by SES. While this raised suspicion that racial differences in school quality were responsible for gap widening, researchers using different methods arrived at opposite conclusions on this question (Fryer and Levitt 2006; Hanushek and Rivkin 2006). Furthermore, the ECLS-K:99 findings regarding the explanatory power of SES did not replicate in other data from the same time period (Murnane et al. 2006).

In this study, I use new nationally representative data from the ECLS-K:2011 (Tourangeau et al. 2012) to replicate and extend past research. While replication studies are crucial to knowledge advancement in the social sciences (Duncan et al. 2012), they remain rare in sociology and education (Makel and Plucker 2014). Without an understanding of whether and when findings replicate, it is difficult to distinguish chance or idiosyncratic findings from consistent educational patterns and to determine how patterns change over time or across contexts. Considering the mixed conclusions from recent studies (Fryer and Levitt 2006; Hanushek and Rivkin 2006; Murnane et al. 2006), new high-quality data provide a valuable opportunity to learn which past results replicate with a new cohort of students.

I extend previous research in several ways. First, I present a more complete picture of how gaps change as students age by considering both differential learning rates by race and changes in test score variance over time. Previous analyses have not teased apart these distinct forces (Fryer and Levitt 2004, 2006). Second, I take advantage of recent methodological advancements to better understand the role that school quality plays in gap changes. Finally, I analyze gaps in working memory (WM). This is important because WM may affect students' learning rates and, therefore, how gaps in academic material develop during schooling. However, little is known about racial gaps in WM.

BACKGROUND

Gap Trends

Black–white test score gaps narrowed considerably during the 1970s and 1980s before progress stagnated in the 1990s. In recent years, gaps have begun to narrow again (Reardon et al. 2015). Most of the evidence on long-term gap trends comes from the NAEP-LTT, which assesses students at ages 9, 13, and 17. Until fairly recently, nationally representative data on gaps in early childhood had not been systematically collected. In 1998, the National Center for Education Statistics (NCES) began the ECLS-K:99, which followed a nationally representative sample of kindergarteners until eighth grade. Using data from this study, Fryer and Levitt (hereafter FL; 2004, 2006) found black–white math and reading gaps at kindergarten entry that were smaller than the gaps in earlier data. FL concluded that these gaps were narrower because of gains made by the recent cohort of black students.

New data from the ECLS-K:2011 study allow researchers to investigate whether these trends have continued. Analyses by Reardon and Portilla (2014) showed that the black–white math and reading gaps at kindergarten entry were approximately .08 standard deviations smaller in 2010 than in 1998, although the difference was not statistically significant for reading and was marginally significant for math. Understanding the sources of these gaps—and how gaps may develop as students age—is crucial to efforts to close the gaps.

Working Memory and Black–White Gaps

Most of the research on black–white gaps has examined literacy and math outcomes; little is known about gaps in executive functioning areas, such as WM. Yet WM deserves attention both because of its importance as an outcome in its own right and because it may affect student learning in other areas. WM is a "domain-general capacity" (Melvy-Lervag and Hulme 2013:271) that involves storing and manipulating information for more complex tasks (Baddeley 2003). Research shows WM predicts fluid intelligence (Redick et al. 2013), reading comprehension, complex learning, reasoning (Engle 2002), and math learning rates (Geary et al. 2012). WM has also been implicated in reading and mathematics difficulties in children (Melvy-Lervag and Hulme 2013). Small-scale studies have shown that students from higher social class backgrounds tend to have stronger WM skills than do students from lower SES backgrounds (Farah et al. 2006) and that childhood poverty predicts lower WM scores in adulthood (Evans and Schamberg 2009).

Black–white WM gaps and gap trends may show different patterns compared to math and reading because WM does not consist of a body of curricular knowledge explicitly taught in schools. For the same reason, comparing changes in math and reading gaps to the WM gap change can provide evidence regarding the role that schools play in gap development. If math and reading gaps grow over the school year but the WM gap does not, this would be consistent with a school quality explanation for the widening of the math and reading gaps. Unlike the ECLS-K:1999, the ECLS-K:2011 includes a WM assessment, enabling these comparisons.

SES and Student Achievement

The association between SES and educational achievement is well known (Coleman et al. 1966; Hedges and Nowell 1999). Because black–white disparities in SES persist as part of the legacy of slavery and racism in the United States (Wilson 2009), many researchers seek to explain racial disparities in students' academic outcomes through racial disparities in SES. While debate continues over how to conceptualize and measure SES, scholars generally agree that the main indicators of SES are income, educational attainment, and occupational prestige (Bradley and Corwin 2002).

Theorists propose that family income affects student learning by providing access to cognitively stimulating experiences and material resources (Chin and Phillips 2004) as well as higher quality health care, which promotes cognitive development in utero and in early childhood (Currie 2005). Additionally, insufficient income often induces parental stress; stressed parents tend to be less warm and supportive toward their children and more punitive, which can negatively influence children's cognitive development (Hackman, Farah, and Meaney 2010; Magnuson and Duncan 2006). Parental education is believed to promote children's cognitive development, because parents with more education tend to make longer utterances to their children and use more complex language (Hoff 2003), hold higher academic expectations, and provide more cognitive stimulation (Davis-Kean 2005; Magnuson and Duncan 2006).

SES and Early Childhood Black–White Test Score Gaps

Research on early childhood black–white gaps has shown variation in the size of gaps and the extent to which SES explains gaps. Table 1 summarizes results from previous large-scale studies. Depending on the sample, year of data collection, and assessment, gaps have ranged from −1.34 standard deviations to −0.40 standard deviations, and various SES measures have explained from 12 to 100 percent of these gaps. Studies using data from the 1980s found the largest early childhood gaps, and gaps of −.25 to −.95 remained after controlling for SES-related covariates.

FL's (2004, 2006) analyses of the nationally representative ECLS-K:1999 yielded results that departed strikingly from past research. FL found narrower gaps at kindergarten entry (math; SD = −.64; reading; SD = −.40), and controlling for a small number of SES-related variables reversed the sign of the reading gap (black students' adjusted advantage; SD = .12) and reduced the math gap by approximately 86 percent (remaining gap; SD = −.09). FL (2004:448) concluded that gains by the recent cohort of black students were "an important part of the explanation" as to why their results diverged from earlier studies.

Murnane and colleagues (2006) tried replicating FL's results using data from the National Institute of Child Health and Human Development's Study of Early Child Care and Youth Development (NICHD-ECCYD). This study included a national—although not nationally representative—sample of students who were similar to those sampled for the ECLS-K:1999 in both age and birth year. However, Murnane and colleagues could not replicate FL's SES finding; like previous

Table 1. Large-scale Studies of Black–White Test Score Gaps in Early Childhood.

					Black–White (*SD*)	
Study	Data	Sample	Controls	Outcome	Unadjusted	Adjusted
Phillips et al. (1998)	Children of the National Longitudinal Survey of Youth	5- to 6-year-olds born 1980 to 1987	Income, mom's education, mom's age at child birth	Peabody Picture Vocabulary Test Revised (PPVT-R)	−1.08	−0.95
	Infant Health and Development Program (IHDP)	5-year-olds, low birth weight, born 1984 to 1985	Income, mom's education, home environment, neighborhood income	PPVT-R	−1.34	−0.55
	IHDP	5-year-olds, low birth weight, born 1984 to 1985	Income, mom's education, home environment, neighborhood income	Wechsler Preschool and Primary Scale of Intelligence	−1.15	−0.25
Fryer and Levitt (2004)	Early Childhood Longitudinal Study–Kindergarten Class of 1998–1999 (ECLS-K):1999	Entering kindergarten fall 1998 (nationally representative)	SES, number of kids' books in home, mom's age at first birth, birth weight, WIC participant	ECLS-K Reading	−0.40	0.12
				ECLS-K Math	−0.64	−0.09
Murnane et al. (2006)	National Institute of Child Health and Human Development	54-month-olds born to healthy mothers age 18 or over (and other restrictions)	SES, maternal sensitivity	Woodcock-Johnson Revised (WJ-R) Math, Applied Problem-Solving subscale	−0.99	−0.42
				WJ-R English	−0.92	−0.37
Yeung and Pfeiffer (2009)	Panel Study of Income Dynamics	Preschool in 1997 (nationally representative)	Model 3: Grandparents' education, teen mom, federal assistance, low birth weight, SES, family structure	WJ-R Math, Applied Problem-Solving	−0.78	Model 3: −0.37
			Model 4: Model 3 plus mother's test score			Model 4: −0.24
				WJ-R English, Letter–Word	−0.43	Model 3: −0.15 Model 4: 0.02

Note: Phillips and colleagues' (1998) and Murnane and colleagues' (2006) standardized gaps were calculated by author from information reported in studies. SES = socioeconomic status; WIC = Women, Infants, and Children Food and Nutrition Services.

researchers, they found large gaps remaining after controlling for SES. The authors hypothesized that the discrepancy between their results and FL's may have been due to differences in the assessments used—the NICHD used more general cognitive assessments, whereas the ECLS-K:1999 focused on academic material taught in schools.

Differential School Quality by Race and Gap Decomposition

While gaps exist before students have experienced much schooling, school segregation and differential school quality by race may widen black–white test score gaps. Schools with high percentages of black or minority students fare worse on a number of school quality measures, such as the fraction of teachers who are novices (Hanushek and Rivkin 2009), teachers' academic preparation, teacher certification status (Lankford, Loeb, and Wyckoff 2002), teacher turnover, and the poverty level of the student body (FL 2004). Black–white disparities on such indicators may persist even after controlling for social class (Condron 2009). Additionally, research shows negative peer effects for black students' academic outcomes in classes with higher percentages of other black students, and these negative effects are stronger for black students than for white students (Hoxby 2000).

Since the Coleman Report (Coleman et al. 1966), researchers have applied test score decompositions to study the effects of school quality on achievement gaps (Hanushek and Rivkin 2006). Although observational data preclude causal inferences, decompositions can provide useful evidence for the following reason: if between-school differences in quality have important effects on student achievement and achievement gaps, between-school differences should explain a large proportion of achievement variation and achievement gaps. Based on their decompositions, Coleman and colleagues (1966:22) concluded that after controlling for SES, "differences between schools account for only a small fraction of differences in pupil achievement," although they also concluded that school quality had stronger effects on minority students' achievement than on white students' achievement. More recent studies decomposing black–white gaps provide evidence that between-school differences and racial segregation have strongly influenced black–white test score gaps (e.g., Page, Murnane, and Willett 2008).

At the center of recent debates over the role of school quality in the development of black–white gaps is a disagreement over the appropriate method for decomposing gaps. Using data from the ECLS-K:1999 and applying a school fixed-effect decomposition strategy, FL (2006) concluded that most of the black–white gap widening from kindergarten to third grade occurred within schools, and therefore differential school quality by race did not explain the widening. Hanushek and Rivkin (hereafter HR; 2006) argued that the FL decomposition strategy was inappropriate because it did not account for segregation. Using the same data, HR applied a decomposition in which each school's contribution to the overall gap was weighted with consideration to the relative number of students from each racial group attending that school. This method led HR to conclude, in contrast to FL, that the majority of gap widening occurred between schools and, therefore, that differential school quality by race played an important role in gap widening.

Reardon (2008) advanced the debate by developing a general decomposition strategy that demonstrated the mathematical relationship between the FL and HR decompositions. Reardon's insight involved decomposing the gap into three—rather than two—portions. Reardon identified an "unambiguously within school" portion, which both FL and HR classify as a within-school gap; an "unambiguously between school" portion, which both FL and HR classify as a between-school gap; and a third "ambiguous" portion that explains the disagreement between the FL and HR decompositions. FL attribute this portion to the within-school gap, whereas HR attribute it to the between-school gap.

Reardon (2008) described convincing reasons as to why the ambiguous portion should be considered part of the between-school gap (see Page et al. 2008; Reardon 2008). Assuming causal interpretations of the decomposition elements, the ambiguous portion could be closed by (1) inducing proportionally greater gains in school mean achievement for schools with higher percentages of black students or (2) eliminating segregation and equalizing outcomes for black and white students within the same schools (Page et al. 2008). The first is clearly a between-school process, and the second is not feasible due to court decisions (Page et al. 2008). This logic therefore leads to the HR gap decomposition, which suggests that efforts to close the black–white gap should focus on between-school forces.

Other Racial/Ethnic Gaps

While the black–white test score gap is perhaps the most studied racial/ethnic gap due to its magnitude and unique historical antecedents, of course the academic achievement of all groups is of interest. National studies do not always include large enough subgroup sample sizes to accurately report the mean achievement of all racial/ethnic groups, but Hispanic–white and Asian–white gaps are often estimated. The ECLS-K:1999 offers some of the best evidence on how Hispanic–white and Asian–white test score gaps develop in early childhood (Reardon et al. 2015). Using these data, FL (2004) reported a Hispanic–white reading gap at kindergarten entry of −.72 standard deviations (adjusted = −.20) and a math gap of −.43 (adjusted = −.06). Unlike black–white gaps, adjusted and unadjusted Hispanic–white gaps narrowed during early schooling (FL 2004). English-fluent Asian students began kindergarten in 1998 scoring higher than white students in math (unadjusted = .15, adjusted = .27) and reading (unadjusted = .34, adjusted = .41), lost some of this advantage in elementary school and regained advantage in middle school (Reardon and Galindo 2009). While in the present study I focus on black–white gaps, I follow previous research (FL 2004, 2006; Reardon and Galindo 2009) by using data from all students. My models therefore also estimate gaps between the other sampled subgroups and white students. I describe these findings in the results section, but future research should examine these gaps more fully.

Summary and Research Questions

Overall, black–white test score gaps have narrowed since the 1970s, and gaps at kindergarten entry may have narrowed between 1998 and 2010. Nonetheless, large math and reading gaps remain. Little is known about black–white WM gaps, yet WM seems to play an important role in academic achievement. Studies vary in the extent to which they find that SES explains early childhood gaps, and FL's unique findings on the explanatory power of SES need replication. While gaps widen over elementary school, researchers have debated what role, if any, school quality plays in this widening.

In this study, I use data from the ECLS-K:2011 to replicate and extend previous research on early childhood black–white test score gaps. Specifically, I ask the following:

Research Question 1: In the fall of kindergarten, what are the black–white gaps in math, reading, and working memory?

Research Question 2: Do these gaps change over kindergarten?

Research Question 3: To what extent does SES explain black–white gaps at kindergarten entry?

Research Question 4: What role does SES play in the development of black–white gaps over kindergarten?

Research Question 5: What role do schools play in the development of black–white gaps over kindergarten?

DATA AND METHODS

The ECLS-K:2011 is an ongoing study conducted by the NCES that is similar to its predecessor, the ECLS-K:1999. The ECLS-K:2011 followed a three-stage sampling design in which (1) 90 primary sampling units (PSUs) of U.S. counties (or groups of contiguous counties) were sampled, (2) public and private schools serving students of kindergarten age were sampled from each PSU (968 schools total), and (3) a target number of 23 kindergarteners were sampled from each selected school (more than 18,000 students total). In the first two stages, units were selected with probability proportional to population size, accounting for planned oversampling of Asians, Native Hawaiians, and other Pacific Islanders. When sampling weights are used, the ECLS-K:2011 is nationally representative of students attending kindergarten during the 2010-to-2011 school year.

Data collection included a parent interview conducted by phone, teacher and principal questionnaires, and direct child cognitive assessments in math, reading, and executive functioning. My analyses use the first wave of data released by the NCES, which contains data collected in the fall and spring of students' kindergarten year (for more information on the ECLS-K:2011, see Tourangeau et al. 2012).

Assessments

Assessments were administered one-on-one by trained child assessors in the fall and spring. Test dates varied by school, with fall assessments occurring between August and December, and spring assessments occurring between January and July. For math and reading, students first answered a set of routing items to determine the appropriate difficulty level of their test questions. Assessments were administered in English or Spanish; students speaking other languages did not participate. For more information on test content, see Tourangeau and colleagues (2012).

I use the theta test metric for math and reading. Theta scores were estimated using an item response theory model, which aims to express all scores across item sets and test waves on a common scale (Tourangeau et al. 2012). Theta reliabilities were .95 for fall and spring reading, and .92 and .94 for fall and spring math, respectively.

As a measure of working memory, students took the Numbers Reversed subtest of the Woodcock-Johnson III Tests of Cognitive Abilities. In this task, the assessor reads increasingly longer series of numbers to the child, who must repeat the numbers in reverse order. This test has a median split-half reliability of .87 (Schrank, McGrew, and Woodcock 2001). I use the standard score metric, which is normed to the child's age. Scoring procedures do not provide standard scores for English-speaking children 62 months or younger with raw scores of 0 or 1; in the weighted sample, this amounts to 9 percent of black test takers and 5 percent of white test takers. WM gaps reported here therefore generalize only to the population of students who obtain valid scores and likely underestimate gaps for the broader population.

Analytic Plan

Estimating Gaps. For each outcome and test occasion separately, I estimate racial/ethnic gaps for the population of kindergarteners in 2010 to 2011 within a regression framework:

$$\begin{aligned} Y_{ita} &= \frac{\mathrm{TESTSCORE}_{ita} - \overline{\mathrm{TESTSCORE}_{ta}}}{SD_{\mathrm{overall}_{ta}}} \\ &= \beta_0 + \beta_{ta}^{(B)}\mathrm{BLACK}_i + \beta_{ta}^{(H)}\mathrm{HISPANIC}_i \\ &\quad + \beta_{ta}^{(A)}\mathrm{ASIAN}_i + \beta_{ta}^{(OR)}\mathrm{OTHER\ RACE}_i + \varepsilon_{ita}, \end{aligned} \quad (1)$$

where i indexes students, t indexes test wave, a indexes the assessment, and ϵ_{ita} is the error term. Notice that in these models, test scores are standardized to mean 0 and standard deviation 1 at each test wave. I include a vector of race/ethnicity dummies (non-Hispanic white being the omitted category) indicating whether a student's parent identified the student as non-Hispanic black, Hispanic (race specified or unspecified), Asian, or "other race" (due to the small number of students identified as multiracial, American Indian/Alaska Native, or Hawaiian/Pacific Islander, I follow FL by combining these students into the category "other race"). The coefficient on each race/ethnicity indicator represents the standardized gap between that group and non-Hispanic whites (also known as G_{ES}, where subscript indicates effect size; Ho 2009). All models incorporate sampling weights and adjust standard errors for the sampling design.[1]

Describing Gap Development over Kindergarten. Several options exist for comparing test score gaps over time (Ho 2009). While different approaches can sometimes yield seemingly contradictory results, such results are often better thought of as different answers to different questions rather than discrepancies (Castellano and Ho 2013; Quinn 2015). It is therefore important to carefully articulate the question that a particular model answers and how that question may differ from questions answered by alternative models.

Changes in standardized gaps (ΔG_{ES}). Perhaps the most common way of examining gaps over time is to compare standardized gaps where the gap at each time point is standardized by the standard deviation for that time point. This is the approach taken by FL, and I also present results using this method:

$$\begin{aligned} &\Delta\hat{G}_{\mathrm{ES}_a}^{(B)} = \hat{G}_{\mathrm{ES}_{a2}}^{(B)} - \hat{G}_{\mathrm{ES}_{a1}}^{(B)} = \hat{\beta}_{a2}^{(B)} - \hat{\beta}_{a1}^{(B)} = \\ &\frac{\overline{\mathrm{TESTSCORE}_{\mathrm{black2}}} - \overline{\mathrm{TESTSCORE}_{\mathrm{white2}}}}{SD_{\mathrm{overall2}}} \\ &- \frac{\overline{\mathrm{TESTSCORE}_{\mathrm{black1}}} - \overline{\mathrm{TESTSCORE}_{\mathrm{white1}}}}{SD_{\mathrm{overall1}}}, \end{aligned} \quad (2)$$

where a indexes assessment, numerical subscripts index test administration, SD_{overall} is the standard deviation for the full sample at the indexed time point, and $\hat{\beta}_{a2}^{(B)}$ and $\hat{\beta}_{a1}^{(B)}$ are the coefficients on the BLACK dummy variable (or analogous coefficients for other groups) for the spring and fall models, respectively (estimated from equation [1]). The $\Delta\hat{G}_{\mathrm{ES}}$ statistic answers the question, "Is the standardized spring gap different from the standardized fall gap?" A negatively signed

$\Delta\hat{G}_{ES}^{(B)}$ indicates that black students have lost relative ground to whites over time, whereas a positively signed $\Delta\hat{G}_{ES}^{(B)}$ indicates that black students have gained relative ground (see Online Appendix A for the $\Delta\hat{G}_{\text{ES}}$ standard error formula).

$\Delta\hat{G}_{\text{ES}}$ is helpful for examining the change in relative group differences over time or the change in the overlap of groups' distributions over time. Relative group differences and distributional overlap are important measures of inequality. Note, however, that $\Delta\hat{G}_{\text{ES}}$ is a function of both the change in mean difference as well as the change in variance. A widening gap over kindergarten as measured by $\Delta\hat{G}_{ES}^{(B)}$ could therefore mean that the average white student made more academic growth than the average black student or that black and white students made equivalent growth while test score variance shrank (or something else).

Differences in change scores. To gain a more complete picture of gap dynamics, I also examine racial differences in academic growth, or fall-to-spring change scores:

$$\begin{aligned}&\Delta\text{SCORE}_{\text{ES}_{ia}} = \\ &\frac{\text{TESTSCORE}_{2ia} - \text{TESTSCORE}_{1ia}}{SD_{\text{overall}\,1a}} = \\ &\beta_0 + \beta_a^{(B)}\text{BLACK}_i + \beta_a^{(H)}\text{HISPANIC}_i + \\ &\quad \beta_a^{(A)}\text{ASIAN}_i + \beta_a^{(OR)}\text{OTHER RACE}_i + \varepsilon_{ia}, \qquad (3)\end{aligned}$$

where terms are as defined earlier (note that test scores are not mean-centered here; dividing by the fall standard deviation simply re-expresses theta growth in a more familiar metric). This model answers the question, "Do black and white students show differential test score growth, on average, over kindergarten?"

To support the academic growth interpretation, change score models require vertically equated test scales. The math and reading theta scores meet this requirement; the WM standard scores, while designed to have interval properties at each test administration, are not vertically equated over time. I therefore do not fit change score models for the WM assessment.

Explaining Gaps and Gap Changes.

To address my research question, "To what extent does SES explain black–white test score gaps at kindergarten entry?" I add a vector of control variables to the model represented by equation (1). To estimate SES-adjusted gap changes (ΔG_{ES}) and SES-adjusted differences in change scores, I refit the models represented by equations (1), (2), and (3) with controls.

Control variables. I use a set of basic controls and a set of student-level SES-related controls similar to those used by FL. The basic controls include the number of months of school a student experienced before testing (as well as the number of months between tests when estimating spring gaps), child's age (in months) at kindergarten entry, an indicator for whether the child is a first-time kindergartener, and an indicator for whether the child is male. To measure SES, I include the continuous SES composite created by NCES, which combines family income, parental occupational prestige score, and parental education. This composite is formed by averaging the standardized (mean 0, standard deviation 1) components (details in Online Appendix B; see Online Appendix C for tables showing interactions between the SES composite and each race/ethnicity indicator[2]). I use the SES composite to make my results comparable to those of FL and other researchers, but I also present tables in Online Appendix D that control for each SES component individually (results are qualitatively similar). Additionally, I include variables indicating whether the child receives support from Women, Infants, and Children Food and Nutrition Services (WIC), whether the mother received WIC support while pregnant, and the number of children's books in the home (divided by 100). Like FL, I include indicators for whether the mother was a teenager or over age 30 at her first child's birth.

Decomposing Gaps and Gap Changes.

For my decompositions, I draw from Reardon (2008), who showed that the unadjusted black–white gap can be decomposed as follows (note that equation [4] is not a model for estimating parameters but an identity expressing the relationship among parameters estimated elsewhere):

$$\hat{G}_{\text{ES}_t}^{(B)} = \hat{\beta}_{1t}\left(1 - \hat{V}\right) + \hat{\beta}_{1t}\hat{V} + \hat{\beta}_{2t}\hat{V}. \qquad (4)$$

In this equation, $\hat{G}_{ES_t}^{(B)}$ is the estimated total gap at time t (estimated from the model represented by equation [1]), and $\hat{V}$ is the estimated variance ratio index of segregation, which can be expressed as $\hat{V} = \bar{p}_s^b - \bar{p}_s^w$, where $\bar{p}_s^b$ is the mean, for black students, of the proportion of sampled students in their school who are black (of the black–white sample), and $\bar{p}_s^w$ is the mean school proportion black (of the black–white sample) for white students. $\hat{\beta}_{1t}(1 - \hat{V})$ represents the unambiguously within-school gap (where $\hat{\beta}_{1t}$ equals FL's within-school gap estimated from a school fixed-effect model), $\hat{\beta}_{2t}\hat{V}$ represents the unambiguously between-school gap (where $\hat{\beta}_{2t}$ equals the coefficient for school proportion black when regressing the student-level outcome on school proportion black and an individual race dummy), and $\hat{\beta}_{1t}\hat{V}$ is the ambiguous portion of the gap (because I decompose adjusted gaps, my decompositions differ slightly; see Online Appendix E).

After decomposing the adjusted fall and spring gaps, I decompose the changes in adjusted gaps by subtracting each component's fall value from its spring value. I use only black and white students who did not switch schools during kindergarten (results are nearly identical when including school switchers).

In all analyses, I use multiply imputed data for missing control variables (after dropping students with missing race or gender data, n = 155), imputed using chained equations with 10 iterations. For each outcome, I include only students with both fall and spring scores.

RESULTS

In Table 2, I present weighted descriptive statistics by race/ethnicity for students with nonmissing data for a given variable. As seen, the number of months of school experienced before and between each assessment was similar across groups.

Research Questions 1 and 2: Describing Gaps and Gap Changes over Kindergarten

G_{ES} and ΔG_{ES}. In Table 3, I present unadjusted standardized gap estimates for the fall and spring of kindergarten ($\hat{G}_{ES}$)—as well as estimates of the changes in gaps ($\Delta\hat{G}_{ES}$)—for reading, math, and WM.

In reading, black students scored approximately .32 standard deviations lower than white students in the fall of kindergarten (column 1), and this gap was approximately .06 standard deviations larger in the spring (column 2). The black–white fall math gap, at −.54 (column 4), also widened over kindergarten, to −.60 (column 5). In contrast, the fall black–white WM gap of −.52 (column 7) did not change significantly (column 9).[3] Again, because valid standard scores are not available for young, low-scoring students, this may underestimate the WM gap.

These patterns differ from the gap patterns for other groups. The Hispanic–white fall reading gap of −0.57 did not show significant change over kindergarten, and the Hispanic–white math gap narrowed from −0.70 to −0.58. In the fall, Asian students scored higher than white students in reading (0.24) and math (0.19), but they lost some of their advantage by spring. No WM gap changed over the school year.

Change Score Differences. In Table 4, I present racial/ethnic differences in test score growth—or change scores—for reading and math. The outcomes in these models are students' spring scores minus fall scores, divided by the fall standard deviation (see equation 3). These models test whether each named racial/ethnic group made equal growth compared to whites over kindergarten (as opposed to testing whether the standardized gaps changed, as in Table 3).

Black and white students made statistically equivalent unadjusted growth in reading (column 1), and black students may have made more math growth than white students (.04 sd, $p < .10$; column 2). Therefore, while it is not the case that kindergarten accelerated white students further ahead of black students, schools may have tended to pull black and white students closer to their respective group's mean, thereby enlarging the average black–white difference relative to the overall standard deviation (models using only black and white students and pooled standard deviation show the same results; see also the shrinking theta variances from fall to spring in Table 2). In contrast, Hispanic students demonstrated more growth than white students in reading and math, and Asian students demonstrated less growth than white students in both subjects.

To summarize the findings from my first two research questions, unadjusted black–white gaps existed at kindergarten entry in 2010 in math, reading, and WM. Math and reading gaps widened by approximately .06 standard deviations over kindergarten, but the WM gap (for students with valid scores) remained constant. Wave-

Table 2. Weighted Descriptive Statistics by Race/Ethnicity.

	White			Black			Hispanic			Asian			Other Race		
Variable	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*
Math theta, fall	−0.29	0.829	7590	−0.78	0.857	2110	−0.93	0.981	3750	−0.10	0.838	1140	−0.50	0.972	930
Math theta, spring	0.598	0.708	7360	0.131	0.773	2000	0.15	0.794	3680	0.682	0.713	1140	0.467	0.723	890
Math theta, fall (standardized)	0.218	0.895	7590	−0.31	0.926	2110	−0.47	1.059	3750	0.423	0.905	1140	−0.01	1.049	930
Math theta, spring (standardized)	0.20	0.924	7360	−0.41	1.008	2000	−0.38	1.036	3680	0.31	0.93	1140	0.03	0.944	890
Reading theta, fall	−0.41	0.83	7620	−0.69	0.789	2120	−0.91	0.863	3740	−0.20	0.984	1180	−0.55	0.88	940
Reading theta, spring	0.635	0.678	7360	0.328	0.764	2000	0.209	0.852	3680	0.71	0.826	1160	0.508	0.761	890
Reading theta, fall (standardized)	0.165	0.945	7620	−0.16	0.898	2120	−0.40	0.983	3740	0.407	1.12	1180	0.006	1.002	940
Reading theta, spring (standardized)	0.183	0.868	7360	−0.21	0.978	2000	−0.36	1.09	3680	0.28	1.058	1160	0.021	0.974	890
Working memory standard score, fall	96.23	16.16	7200	87.83	15.65	1920	87.61	15.20	3370	99.53	17.69	1030	94.89	16.63	870
Working memory standard score, spring	98.01	16.16	7350	89.54	17.15	1990	90.07	16.84	3680	100.4	17.41	1140	96.68	16.54	890
Working memory standard score, fall (standardized)	0.17	0.974	7200	−0.34	0.943	1920	−0.35	0.916	3370	0.369	1.066	1030	0.09	1.002	870
Working memory standard score, spring (standardized)	0.176	0.943	7360	−0.32	1.00	1990	−0.29	0.983	3680	0.317	1.016	1140	0.099	0.965	890
WIC	0.349	0.477	6980	0.796	0.403	1770	0.744	0.437	3260	0.279	0.449	980	0.579	0.494	860
WIC (while mother pregnant)	0.307	0.461	6970	0.74	0.439	1760	0.691	0.462	3250	0.239	0.427	970	0.525	0.50	850
Male	0.517	0.50	7640	0.511	0.50	2130	0.52	0.50	3790	0.464	0.499	1200	0.516	0.50	940
First-time kindergartner	0.953	0.212	7580	0.917	0.276	2100	0.944	0.229	3710	0.959	0.199	1130	0.948	0.222	940
Mother >30	0.189	0.392	6600	0.066	0.248	1590	0.078	0.269	2810	0.235	0.424	770	0.146	0.353	790
Teen mother	0.176	0.381	6600	0.444	0.497	1590	0.38	0.485	2810	0.064	0.245	770	0.317	0.466	790
Age at kindergarten entry (months)	66.64	4.417	7580	65.7	5.14	2100	65.42	4.592	3710	64.59	4.542	1130	66.45	4.313	930
Number of children's books	112.4	152.0	6740	44.21	56.02	1690	51.31	121.9	2900	63.98	104	840	90.34	120.5	820
SES	0.18	0.736	7100	−0.38	0.7	1840	−0.53	0.712	3330	0.379	0.899	1030	−0.04	0.782	880
School % nonwhite	26.3	23.2	7480	73.1	27	2050	70.5	29	3750	61.4	29.8	1190	46.9	27.5	910
Months school, fall	1.894	0.957	7380	1.886	0.886	1990	1.981	0.983	3610	1.99	1.025	1170	1.803	0.942	890
Months school, spring	7.848	0.884	7310	7.878	0.817	1960	7.878	0.912	3580	7.717	0.928	1150	7.816	0.875	880
Months school, between	5.961	0.95	7370	6.002	0.909	2000	5.887	1.032	3690	5.747	0.989	1160	6.017	0.933	890

Note: Sample sizes rounded to nearest 10, per National Center for Education Statistics requirements. Samples include complete cases for each variable. Sampling weight = W1C0. WIC = receives support through Women, Infant, and Children Food and Nutrition Services; mother >30 = mother age 30 or older when first child born; teen mother = mother under age 20 at first child's birth; SES = composite of parental income, occupational prestige, and education; months school = number of months of school before test for season indicated (*between* indicates number of months of school between fall and spring tests).

Table 3. Kindergarten Fall and Spring Gaps, and Changes in Gaps, for Reading, Math, and Working Memory.

	(1)	(2) Reading	(3)	(4)	(5) Math	(6)	(7)	(8) Working Memory	(9)
	Fall b (SE)	Spring b (SE)	ΔG_{ES} (Spring–Fall)	Fall b (SE)	Spring b (SE)	ΔG_{ES} (Spring–Fall)	Fall b (SE)	Spring b (SE)	ΔG_{ES} (Spring–Fall)
Black	−.318***	−.380***	−.062*	−.542***	−.603***	−.061*	−.517***	−.502***	.015
	(.045)	(.046)	(.031)	(.042)	(.047)	(.028)	(.050)	(.050)	(.045)
Hispanic	−.569***	−.532***	.037	−.695***	−.581***	.114***	−.525***	−.479***	.046
	(.043)	(.038)	(.028)	(.046)	(.043)	(.027)	(.029)	(.028)	(.025)
Asian	.236***	.116*	−.120**	.194***	.129**	−.065*	.187***	.189***	.002
	(.054)	(.047)	(.035)	(.044)	(.043)	(.027)	(.050)	(.049)	(.044)
Other race	−.142*	−.151**	−.009	−.209**	−.173**	.036	−.077	−.062	.015
	(.068)	(.054)	(.044)	(.073)	(.057)	(.043)	(.049)	(.045)	(.042)
Constant	.179***	.186***		.236***	.208***		.184***	.198***	
	(.027)	(.027)		(.028)	(.036)		(.026)	(.025)	
Approximate N	14,970	14,970	14,970	14,890	14,890	14,890	13,830	13,830	13,830

Note: Standard errors account for sampling design. Sample sizes rounded to nearest 10, per National Center for Educational Statistics requirements. Outcomes standardized to SD = 1 at each test occasion. Omitted group = non-Hispanic whites. Results obtained by analyzing multiply imputed data (imputed using chained equations with 10 iterations); therefore R^2 is not estimated. Sampling weight = W1C0.

†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

Table 4. Racial/Ethnic Differences in Reading and Math Change Scores over Kindergarten.

	(1) Reading *b* (*SE*)	(2) Math *b* (*SE*)
Black	−0.020 (0.032)	0.043† (0.022)
Hispanic	0.096* (0.041)	0.215*** (0.028)
Asian	−0.132*** (0.034)	−0.088** (0.026)
Other race	0.008 (0.033)	0.065† (0.039)
Constant	1.177*** (0.021)	0.948*** (0.016)
Approximate *N*	14,970	14,890

Note: Standard errors account for sampling design. Sample sizes rounded to nearest 10, per National Center for Education Statistics requirements. Outcomes = students' spring–fall change scores, divided by fall standard deviation. Omitted group = non-Hispanic whites. Results obtained by analyzing multiply imputed data (imputed using chained equations with 10 iterations); therefore, R^2 not estimated. Sampling weight = W1C0.
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

standardized black–white math and reading gaps widened, not because black students learned less than white students but because test score variance shrank from fall to spring.

Explaining Gaps and Gap Changes over Kindergarten

Research Question 3: Does SES Explain Fall Gaps? As column 1 of Table 5 shows, the SES-related controls reverse the sign of the fall black–white reading gap, giving black students a significant adjusted advantage of approximately .08 standard deviations. The controls explain about 75 percent of the black–white math gap (reduced from −.54 to −.14, column 4) but only 54 percent of the fall WM gap (reduced from −.52 to −.24; column 7).

The same controls dramatically reduce the Hispanic–white fall reading (–.57 to −.13) and math (–.70 to −.23) gaps but reduce the WM gap less (from −.53 to −.26). While the SES controls have little effect on Asian students' fall math and reading advantage, controlling for background reduces Asian students' WM advantage to nonsignificance.

Research Question 4: Does SES Explain Gap Changes over Kindergarten? If SES explained the black–white math and reading gap widening over kindergarten, one would expect to see SES-adjusted spring gaps of similar magnitude to SES-adjusted fall gaps. However, I find that black students lost their adjusted reading advantage by spring, with the adjusted gap going from .08 (column 1 of Table 5) to −.05 (column 2), and this loss was statistically significant (column 3). The adjusted black–white math gap widened significantly, by .10 standard deviations (column 6).

In contrast, the adjusted Hispanic–white reading gap did not change over kindergarten, and the adjusted Hispanic–white math gap narrowed somewhat. In both math and reading, Asian students lost some of their adjusted advantage. None of the adjusted WM gaps changed.

SES and change score differences. In Table 6, I present racial/ethnic differences in change scores (divided by the fall standard deviation), adjusting for the full set of controls from the spring models in Table 5 (controls not shown to conserve space). In contrast to Table 4, which shows black and white students making equal unadjusted growth over kindergarten, Table 6 shows that black students made less adjusted growth than white students in math and reading.

While Hispanic students' unadjusted reading gain scores were significantly higher than white students' (Table 4), Table 6 shows that no adjusted mean difference existed between these groups. After adding controls, Hispanic students still gained significantly more math skill than white students, but the adjusted advantage is approximately half of the unadjusted advantage. With controls, Asian students gained less in math and reading than did white students.

To recapitulate, Research Question 4 investigated whether SES explained the widening of

Table 5. Adjusted Gaps and Changes in Adjusted Gaps over Kindergarten for Reading, Math, and Working Memory.

	(1)	(2) Reading	(3)	(4)	(5) Math	(6)	(7)	(8) Working Memory	(9)
	Fall *b* (*SE*)	Spring *b* (*SE*)	ΔG_{ES} (Spring–Fall)	Fall *b* (*SE*)	Spring *b* (*SE*)	ΔG_{ES} (Spring–Fall)	Fall *b* (*SE*)	Spring *b* (*SE*)	ΔG_{ES} (Spring–Fall)
Black	0.075*	−0.046	−0.121***	−0.136**	−0.237***	−0.101***	−0.243***	−0.255***	−0.012
	(0.033)	(0.039)	(0.025)	(0.041)	(0.046)	(0.027)	(0.032)	(0.043)	(0.035)
Hispanic	−0.134***	−0.137***	−0.003	−0.232***	−0.153***	0.079**	−0.258***	−0.238***	0.020
	(0.031)	(0.034)	(0.023)	(0.038)	(0.044)	(0.026)	(0.027)	(0.034)	(0.028)
Asian	0.229***	0.156***	−0.073**	0.233***	0.182***	−0.051*	0.040	0.055	0.015
	(0.042)	(0.039)	(0.028)	(0.035)	(0.039)	(0.023)	(0.038)	(0.042)	(0.036)
Other race	0.042	−0.009	−0.051†	−0.028	−0.021	0.007	0.030	0.032	0.002
	(0.045)	(0.039)	(0.030)	(0.054)	(0.048)	(0.032)	(0.033)	(0.046)	(0.037)
Male	−0.142***	−0.178***		−0.020	−0.031†		−0.078***	−0.075***	
	(0.015)	(0.016)		(0.017)	(0.017)		(0.016)	(0.018)	
First-time kindergartner	−0.611***	−0.238***		−0.608***	−0.363***		0.849***	0.811***	
	(0.046)	(0.049)		(0.052)	(0.062)		(0.046)	(0.062)	
Age at kindergarten entry (months)	0.035***	0.031***		0.049***	0.043***		−0.058***	−0.049***	
	(0.002)	(0.002)		(0.002)	(0.003)		(0.002)	(0.002)	
Months school	0.232***	0.099***		0.154***	0.065***		0.014	−0.006	
	(0.010)	(0.017)		(0.010)	(0.018)		(0.012)	(0.016)	
SES	0.387***	0.335***		0.373***	0.338***		0.276***	0.253***	
	(0.017)	(0.018)		(0.017)	(0.014)		(0.015)	(0.013)	
SES^2	−0.027*	−0.053***		−0.058***	−0.041**		−0.037**	−0.037**	
	(0.011)	(0.013)		(0.013)	(0.014)		(0.012)	(0.012)	
WIC	−0.091*	−0.096**		−0.077†	−0.081*		−0.095*	−0.023	
	(0.036)	(0.037)		(0.040)	(0.035)		(0.038)	(0.037)	
WIC (while pregnant)	−0.021	0.006		−0.064†	−0.050		−0.026	−0.059	
	(0.034)	(0.038)		(0.034)	(0.035)		(0.035)	(0.039)	
Mother >30	0.155***	0.057*		0.127***	0.086**		0.070*	0.057*	
	(0.026)	(0.026)		(0.024)	(0.026)		(0.027)	(0.026)	
Teen mother	−0.154***	−0.121***		−0.142***	−0.126***		−0.048*	−0.095***	
	(0.025)	(0.025)		(0.024)	(0.024)		(0.022)	(0.023)	
Books/100	0.078***	0.069***		0.078***	0.071***		0.061***	0.042***	
	(0.014)	(0.013)		(0.014)	(0.012)		(0.014)	(0.012)	
Books^2/100^2	−0.003***	−0.003***		−0.003***	−0.003***		−0.002***	−0.002*	

(continued)

Table 5.
(Continued)

	(1)	(2) Reading	(3)	(4)	(5) Math	(6)	(7)	(8) Working Memory	(9)
	Fall *b* (*SE*)	Spring *b* (*SE*)	ΔG_{ES} (Spring–Fall)	Fall *b* (*SE*)	Spring *b* (*SE*)	ΔG_{ES} (Spring–Fall)	Fall *b* (*SE*)	Spring *b* (*SE*)	ΔG_{ES} (Spring–Fall)
	(0.001)	(0.001)		(0.001)	(0.001)		(0.001)	(0.001)	
Months between fall and spring tests		0.011			0.015			0.011	
		(0.015)			(0.014)			(0.013)	
Constant	−2.049***	−2.495***		−2.783***	−2.923***		3.195***	2.709***	
	(0.134)	(0.235)		(0.175)	(0.253)		(0.137)	(0.168)	
Approximate *N*	14,970	14,970	14,970	14,890	14,890	14,890	13,830	13,830	13,830

Note: Standard errors account for sampling design. Results obtained by analyzing multiply imputed data (imputed using chained equations with 10 iterations); therefore R^2 not estimated. Sampling weight = W1C0. Sample sizes rounded to nearest 10, per National Center for Education Statistics requirements. Outcomes standardized to *SD* = 1 at each test occasion. Omitted group = non-Hispanic whites. Months school = number of months of school before test. SES = composite of parental income, occupational prestige, and education. WIC = receives support through Women, Infant, and Children Food and Nutrition Services; mother >30 = mother age 30 or older when first child born; teen mother = mother under age 20 at first child's birth; books = number of children's books in home (in 100s).
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

Table 6. Adjusted Racial/Ethnic Differences in Reading and Math Change Scores over Kindergarten.

	(1) Reading *b* (*SE*)	(2) Math *b* (*SE*)
Black	−.116*** (.029)	−.060** (.021)
Hispanic	.013 (.032)	.108*** (.022)
Asian	−.088** (.029)	−.078*** (.022)
Other race	−.050† (.026)	.011 (.032)
Full set of controls?	Yes	Yes
Approximate *N*	14,970	14,890

Note: Standard errors account for sampling design. *Full set of controls* refers to controls from spring models. Results obtained by analyzing multiply imputed data (imputed using chained equations with 10 iterations); therefore, R^2 is not estimated. Sampling weight = W1C0. Outcomes = spring–fall change scores divided by fall whole-sample standard deviation. Omitted group = non-Hispanic whites. Sample sizes rounded to nearest 10, per National Center for Education Statistics requirements.
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

the black–white math and reading gaps. SES did not explain the widening gaps; in fact, SES-adjusted gaps widened more than unadjusted gaps. Adjusted math and reading gaps widened primarily because black students learned less math and reading over kindergarten than did white students from similar backgrounds.

Research Question 5: Do Schools Contribute to Black–White Gap Widening?

Decomposing gap changes. In Table 7, I present the results of the decompositions for adjusted reading and math gaps and gap changes (adjusted gap estimates differ slightly from those in previous tables due to sample restrictions described earlier). For reading, after combining the adjusted ambiguous portion of gap change with the adjusted unambiguously between-school portion (as argued for earlier), approximately 71 percent of black students' loss of adjusted ground occurred between schools (or 76 percent when decomposing unadjusted reading gap change). This coheres with a school quality explanation for reading gap widening.

In math, the decomposition looks different. First, while the overall adjusted math gap change is negatively signed (indicating a widening gap), the adjusted unambiguously between-school portion of the adjusted gap change is positively signed (but offset by negatively signed changes for other portions). This is because school proportion black negatively predicts students' math scores in the fall but positively predicts students' math scores in the spring (controlling for background variables). Combining the adjusted ambiguous portion of the math gap change with the adjusted unambiguously between-school portion, approximately 63 percent of the adjusted gap widening occurred within schools, and 37 percent occurred between schools (for unadjusted gaps, 35 percent of gap widening occurred between schools).[4]

In summary, the gap decomposition results are consistent with differential school quality by race playing a large role in black students' loss of ground in reading but much less of a role in the widening of the math gap.

Alternative Explanations

The results presented here are consistent with a story in which black–white reading and math gaps at kindergarten entry are largely a function of SES, but after kindergarten entry, differential school quality by race widens the reading gap. However, the reading gap may have widened due to unobserved nonschool factors correlated with school percentage black. In Online Appendix F, I use evidence to argue against alternative explanations. For example, school proportion black positively and significantly predicts SES-adjusted fall reading scores (as recovered from Table 7). By spring, this effect is dramatically reduced and marginally significant. This is more consistent with a school quality explanation for gap change.

Table 7. Decomposing Adjusted Black–White Reading and Math Gaps and Gap Changes over Kindergarten.

	Reading				Math			
	Total Adjusted Gap	Adjusted Within-School	Adjusted Ambiguous	Adjusted Between-School	Total Adjusted Gap	Adjusted Within-School	Adjusted Ambiguous	Adjusted Between-School
Fall	0.066	−0.014	−0.016	0.096	−0.144	−0.064	−0.075	−0.005
%		−21%	−24%	145%		44%	53%	3%
Spring	−0.054	−0.048	−0.057	0.052	−0.225	−0.115	−0.137	0.027
%		90%	106%	−96%		51%	61%	−12%
Spring–fall	−0.12	−0.034	−0.041	−0.044	−0.081	−0.051	−0.062	0.032
%		29%	34%	37%		63%	77%	−40%

Note: Analyses use only the black–white sample. Gap components may not sum to total due to rounding; percentages were calculated prior to rounding. Gaps adjusted for full set of controls. Math: black n = 1,890; white n = 7,190. Reading: black n = 1,900; white n = 7,210. Results obtained by analyzing multiply imputed data (imputed using chained equations with 10 iterations). Sampling weight = W1C0. Sample sizes rounded to nearest 10, per National Center for Education Statistics requirements.

Additionally, I tested whether the gap widening in math could be explained by (1) racial differences in school readiness or prekindergarten experiences, (2) racial differences in parental beliefs about the importance of math, (3) differential math learning rates stemming from the fall working memory gap, (4) student–teacher racial mismatch, or (5) black students being assigned to less effective teachers than white students within the same schools. These exploratory analyses did not produce convincing supporting evidence; however, data for these analyses are limited and further research is needed. See Online Appendix G for a full description of the analyses and results.

DISCUSSION

In this study, I use new data from the ECLS-K: 2011 to replicate and extend previous research on early childhood black–white math and reading gaps and to shine new light on black–white working memory gaps. At kindergarten entry in 2010, black–white gaps existed in reading (–.32 standard deviations, equivalent to the difference between the 50th and 37th percentiles of a normal distribution), math (–.54 standard deviations, or 50th percentile to 29th), and working memory (.52 standard deviations, or 50th percentile to 30th; again, these data may underestimate the WM gap given that young, low-scoring students do not have scores). Unadjusted math and reading gaps widened by approximately .06 standard deviations over kindergarten, but the WM gap did not change. Unadjusted math and reading gaps widened not because white students learned more than black students but because test score variance shrank. Student SES explained all of the fall reading gap, 75 percent of the fall math gap, and about 54 percent of the fall WM gap. SES could not explain why math and reading gaps widened over the school year; in fact, controlling for background, black students learned less math and reading than did white students.

Gap decompositions and follow-up analyses provide evidence that differential school quality by race played a large role in black students' loss of ground to white students in reading. In contrast, the majority of math gap widening occurred within schools, suggesting that differential school quality explained much less of the math gap widening. Several other hypothesized explanations for the widening math gap were not supported by the data.

SES and Black–White Gaps

The finding that black students scored significantly higher than white students in reading at kindergarten entry after controlling for SES replicates FL's (2004) result from the ECLS-K:1999. This is consistent with literature on the "net black advantage" (Bennett and Xie 2003) showing that black students tend to attain higher levels of education compared to white students with similar SES and school records (Merolla 2013). Some evidence suggests that the net black advantage may be in part due to parental aspirations, where the historically limited access to educational opportunities for African Americans in the United States has led black families to place increased importance on education (Merolla 2013). Higher educational aspirations held by black parents compared to similar-SES white parents could help explain black students' adjusted reading advantage at kindergarten entry.

Why might SES explain less of the WM gap than the reading and math gaps? One potential explanation is that racial gaps in WM are driven by environmental factors that are not well captured by the imperfect SES measures in this study but that affect general cognitive development during the "sensitive period" (Knudsen 2004) of early childhood—factors such as external stimuli, stress, and nutrition (Hackman et al. 2010). There is growing agreement that controlling for SES does not equalize conditions for black and white students (Downey 2008), and black students may be worse off than white students in areas affecting WM, even after controlling for the SES measures available in the ECLS-K. For example, evidence suggests that the effect of childhood poverty on adult WM is mediated by childhood stress (Evans and Schamberg 2009). Black children's WM scores may therefore be depressed, in part, due to stress not experienced by white students from similar backgrounds. Additionally, longer periods of childhood poverty seem to be more detrimental to children's cognitive development than shorter periods (Magnuson and Duncan 2006), and family wealth explains children's test scores over and above current income (Orr 2003). Given that the ECLS-K measures income at one time point and does not measure wealth, my models may not equalize black and white students on important dimensions of SES affecting WM. Such unmeasured dimensions of SES may affect WM more than math and reading because children's math and reading skills may be more responsive to activities parents commonly engage their children in, such as reading to them and teaching them to count. Such explicit skill cultivation could help compensate for environmental influences in a way not seen for WM. Future research should examine whether the dimensions of SES that are unmeasured here further explain black–white WM gaps.

The Role of Schools. Despite the fact that $\Delta G_{ES}^{(B)}$ shows widening math and reading gaps over kindergarten, one might argue that schools nevertheless serve as "equalizers" because black and white students make statistically equivalent academic growth (before adding controls). After all, the existence of the fall gaps suggests that black students had been learning less than white students prior to kindergarten entry. At the heart of this argument is the contention that observed spring gaps would have been larger had it not been for schools. While this may be true, another counterfactual to consider (as opposed to the "no-school" counterfactual) is the counterfactual in which school quality does not differ by race (Downey, von Hippel, and Broh [2004] make a similar point). This is the counterfactual that the gap decomposition models attempt to simulate when estimating the between-school portion of gaps and gap changes.

Why do between-school differences in school quality seem to affect reading gaps more than math gaps? This result may seem unexpected, given arguments that students' math achievement is more sensitive to schooling than is their reading achievement (Hedges and Nowell 1999). One potential explanation relates to the children's age—in early childhood, schools tend to more strongly emphasize literacy instruction than math instruction (Engel, Claessens, and Finch 2013). Consequently, literacy may be the area in which instructional effectiveness varies most across schools. Given that several hypothesized explanations for within-school math gap widening did not yield convincing supporting evidence (see Online Appendix G), inquiry into explanations that more sharply focus on instructional processes within schools and classrooms, or on how students' experiences prior to kindergarten prepare them for successful math learning, may be a fruitful area of future investigation.

Limitations. As with all observational studies, these results may suffer from bias due to omitted variables. While these analyses provide evidence for a school quality explanation for reading gap widening, we cannot completely rule out the possibility that unobserved nonschool factors are responsible. Furthermore, if we could confirm a causal effect of schools, this would not reveal which specific school-based factors were responsible for widening the gap. Similarly, although SES helps to explain gaps in these models, these analyses cannot identify the specific mechanisms responsible for gap formation. These findings are therefore most useful in building an understanding of the broad categories of factors that affect gaps, rather than suggesting specific solutions. An important limitation to the WM analyses is that young, low-scoring, English-speaking students do not have valid standard scores. These gap estimates therefore do not generalize to all students and likely underestimate the size of the black–white WM gap. Finally, we must keep in mind that these results apply to the kindergarten school year. As additional data from the ECLS-K:2011 become available, further research on how these gaps develop should be conducted.

CONCLUSION

The picture that emerges from these analyses is one in which black–white math and reading gaps prior to kindergarten entry are largely functions of SES. SES explains less of the WM gap, however (for students with valid WM scores). After kindergarten entry, math and reading gaps widen, whereas the WM gap is constant. Differential school quality by race may be primarily responsible for the widening reading gap, but most of the math gap widening is yet unexplained.

While these analyses provide important new evidence on black–white gaps in early childhood, additional questions emerge. Why do math and reading gaps widen over kindergarten while the WM gap does not? Why does SES explain more of the reading and math gaps than the WM gap? Why does the reading gap widen primarily between schools while the math gap widens within schools? These are important questions for future research.

RESEARCH ETHICS

Data used in this study were collected by the National Center for Education Statistics and distributed under a restricted-use data licensing agreement. The data file contained no personal identifiers, and all Institute of Education Sciences protocols for use and storage of restricted-use data were followed by the author. This presentation of analytic results was approved by the Institute of Education Sciences for dissemination to non-licensed personnel.

ACKNOWLEDGMENTS

I am grateful to John Robert Warren, North Cooc, Celia Gomez, Ken Shores, James Kim, Beth Schueler, and three anonymous reviewers for comments on earlier drafts of this paper and to Richard Murnane, Dan Koretz, Sean Reardon, Andrew Ho, Sara Heller, Catherine Snow, and Joe McIntyre for comments and discussion. All errors and omissions are my own.

FUNDING

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Harvard Graduate School of Education Dean's Summer Fellowship.

SUPPLEMENTAL MATERIAL

The online appendices are available at soe.sagepub.com/supplemental.

NOTES

1. I use sampling weight W1C0 (with multiply imputed data, as described). To adjust standard errors for the sampling design, I use strata and PSU variables with the svy command suite in Stata 12.
2. No interaction was significant for black students; certain other interactions were significant.
3. The working memory scores had a somewhat bimodal fall distribution (less so in spring). Sensitivity analyses using the ordinal effect size gap measure V (Ho 2009) show black–white gaps and gap changes similar to those in Table 3.
4. For math and reading, decompositions of adjusted change score differences show breakdowns similar to the $\Delta\hat{G}_{ES}^{(B)}$ decompositions.

REFERENCES

Baddeley, Alan. 2003. "Working Memory: Looking Back and Looking Forward." *Nature Reviews Neuroscience* 4:829–39.

Bennett, Pamela R. and Yu Xie. 2003. "Revisiting Racial Differences in College Attendance: The Role of Historically Black Colleges and Universities." *American Sociological Review* 68:567–80.

Bradley, Robert H. and Robert F. Corwin. 2002. "Socioeconomic Status and Child Development." *Annual Review of Psychology* 53:371–99.

Castellano, Katherine E. and Andrew D. Ho. 2013. *A Practitioner's Guide to Growth Models*. Washington, DC: Council of Chief State School Officers.

Chin, Tiffani and Meredith Phillips. 2004. "Social Reproduction and Child-rearing Practices: Social Class, Children's Agency, and the Summer Activity Gap." *Sociology of Education* 77:185–210.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert York. 1966. *Equality of Educational Opportunity*. Washington, DC: National Center for Education Statistics.

Condron, Dennis J. 2009. "Social Class, School and Non-school Environments, and Black/White Inequalities in Children's Learning." *American Sociological Review* 74:683–708.

Currie, Janet M. 2005. "Health Disparities and Gaps in School Readiness." *Future of Children* 15:117–38.

Davis-Kean, Pamela E. 2005. "The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment." *Journal of Family Psychology* 19:294–304.

Downey, Douglas B. 2008. "Black/White Differences in School Performance: The Oppositional Culture Explanation." *Annual Review of Sociology* 34:107–26.

Downey, Douglas B., Paul T. von Hippel, and Beckett Broh. 2004. "Are Schools the Great Equalizer? Cognitive Inequality During the Summer Months and the School Year." *American Sociological Review* 69: 613–35.

Duncan, Greg J., Mimi Engel, Amy Claessens, and Chantelle J. Dowsett. 2012. "The Value of Replication for Developmental Science." Working paper. Retrieved February 6, 2015 (http://sites.uci.edu/gduncan/files/2013/06/Replication-paper-single-spaced.pdf).

Engel, Mimi, Amy Claessens, and Maida A. Finch. 2013. "Teaching Students What They Already Know? The (Mis) Alignment between Mathematics Instructional Content and Student Knowledge in Kindergarten. *Educational Evaluation and Policy Analysis* 35: 157–78.

Engle, Randall W. 2002. "Working Memory Capacity as Executive Attention." *Current Directions in Psychological Science* 11:19–23.

Evans, Gary W. and Michelle A. Schamberg. 2009. "Childhood Poverty, Chronic Stress, and Adult Working Memory." *Proceedings of the National Academy of Sciences* 106:6545–49.

Farah, Martha J., David M. Shera, Jessica H. Savage, Laura Betancourt, Joan M. Giannetta, Nancy L. Brodsky, Elsa K. Malmud, and Hallam Hurt. 2006. "Childhood Poverty: Specific Associations with Neurocognitive Development." *Brain Research* 1110:166–74.

Fischer, Claude S. and Michael Hout. 2006. *Century of Difference: How America Changed the Last One Hundred Years*. New York: Russell Sage Foundation.

Fryer, Roland G. and Steven D. Levitt. 2004. "Understanding the Black–White Test Score Gap in the First Two Years of School." *Review of Economics and Statistics* 86:447–64.

Fryer, Roland G. and Steven D. Levitt. 2006. "The Black–White Test Score Gap through Third Grade." *American Law and Economics Review* 8:249–81.

Geary, David C., Mary K. Hoard, Lara Nugent, and Drew H. Bailey. 2012. "Mathematical Cognition Deficits in Children with Learning Disabilities and Persistent Low Achievement: A Five-year Prospective Study." *Journal of Educational Psychology* 104:206–23.

Hackman, Daniel A., Martha J. Farah, and Michael J. Meaney. 2010. "Socioeconomic Status and the Brain: Mechanistic Insights from Human and Animal Research." *Nature Reviews Neuroscience* 11: 651–59.

Hanushek, Eric A. and Steven G. Rivkin. 2006. *School Quality and the Black–White Achievement Gap* (No. w12651). Cambridge, MA: National Bureau of Economic Research.

Hanushek, Eric A. and Steven G. Rivkin. 2009. "Harming the Best: How Schools Affect the Black–White Achievement Gap." *Journal of Policy Analysis and Management* 28:366–93.

Hedges, Larry V. and Amy Nowell. 1999. "Changes in the Black–White Gap in Achievement Test Scores. " *Sociology of Education* 72:111–35.

Ho, Andrew D. 2009. "A Nonparametric Framework for Comparing Trends and Gaps across Tests." *Journal of Educational and Behavioral Statistics* 34: 201–28.

Hoff, Erik. 2003. "The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development through Maternal Speech. " *Child Development* 74:1368–78.

Hoxby, Caroline. 2000. *Peer Effects in the Classroom: Learning from Gender and Race Variation* (No. w7867). Cambridge, MA: National Bureau of Economic Research.

Knudsen, Eric I. 2004. "Sensitive Periods in the Development of the Brain and Behavior." *Journal of Cognitive Neuroscience* 16:1412–25.

Lankford, Hamilton, Susanna Loeb, and James Wyckoff. 2002. "Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis." *Educational Evaluation and Policy Analysis* 24:37–62.

Magnuson, Katherine A. and Greg J. Duncan. 2006. "The Role of Family Socioeconomic Resources in the Black–White Test Score Gap among Young Children." *Developmental Review* 26:365–99.

Makel, Matthew C. and Jonathan A. Plucker. 2014. "Facts Are More Important Than Novelty: Replication in Education Sciences." *Educational Researcher* 46:304–16.

Melvy-Lervag, M. M. and Charles Hulme. 2013. Is Working Memory Training Effective? A Meta-analytic Review. *Developmental Psychology* 49: 270–91. doi:10.1037/a0028228

Merolla, David M. 2013. "The Net Black Advantage in Educational Transitions: An Education Careers Approach." *American Educational Research Journal* 50:895–924.

Murnane, Richard J., John B. Willett, Kristen L. Bub, and Kathleen McCartney. 2006. "Understanding Trends in the Black–White Achievement Gaps during the First Years of School." *Brookings-Wharton Papers on Urban Affairs* 2006:97–135.

Orr, Amy J. 2003. "Black–White Differences in Achievement: The Importance of Wealth." *Sociology of Education* 76:281–304.

Page, Lindsay C., Richard J. Murnane, and John B. Willett. 2008. *Trends in the Black–White Achievement Gap: Clarifying the Meaning of within- and between-school Achievement Gaps* (NBER Working Paper No. 14213). Cambridge, MA: National Bureau for Economic Research.

Phillips, Meredith, Jeanne Brooks-Gunn, Greg J. Duncan, Pamela Klebanov, and Jonathan Crane. 1998. "Family Background, Parenting Practices, and the Black–White Test Score Gap." Pp. 103–45 in *The Black–White Test Score Gap*, edited by C. Jencks and M. Phillips. Washington, DC: Brookings Institution.

Quinn, David M. 2015. "Black-White Summer Learning Gaps: Interpreting the Variability of Estimates Across Representations." *Educational Evaluation and Policy Analysis*. 37:50–69. doi:10.3102/0162373714534522

Reardon, Sean F. 2008. "Thirteen Ways of Looking at the Black–White Test Score Gap." Working paper, Stanford University. Retrieved February 6, 2015 (http://web.stanford.edu/group/irepp/cgi-bin/joomla/working-papers.html)

Reardon, Sean F. and Claudia Galindo. 2009. "The Hispanic–White Achievement Gap in Math and Reading in the Elementary Grades." *American Educational Research Journal* 46:853–91.

Reardon, Sean F. and Ximena Portilla. 2014. "Recent Trends in Socioeconomic and Racial School Readiness Gaps at Kindergarten Entry." Paper presented at the Annual Meeting of the Society for Research on Educational Effectiveness, February 2014, Washington, DC.

Reardon, Sean F., Joseph P. Robinson-Cimpian, and Ericka S. Weathers. 2015. "Patterns and Trends in Racial/Ethnic and Socioeconomic Academic Achievement Gaps." Pp. 491–509 in *Handbook of Research in Education Finance and Policy*, 2nd ed., edited by H. A. Ladd and M. E. Goertz. New York: Routledge.

Redick, Thomas S., Zach Shipstead, Tyler L. Harrison, Kenny L. Hicks, David E. Fried, David Z. Hambrick, Michael J. Krane, and Randall W. Engle. 2013. "No Evidence of Intelligence Improvement after Working Memory Training: A Randomized, Placebo-controlled Study." *Journal of Experimental Psychology: General* 142:359–79.

Schrank, Frederick A., Kevin S. McGrew, and Richard R. Woodcock. 2001. "Technical Abstract (Woodcock-Johnson III Assessment Service Bulletin No. 2)." Itasca, IL: Riverside.

Tourangeau, Karen, Christine Nord, Thanh Lê, Alberto G. Sorongon, Mary C. Hagedorn, Peggy Daly, and Michelle Najarian. 2012. *Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011), User's Manual for the ECLS-K: 2011 Kindergarten Data File and Electronic Codebook* (NCES 2013-061). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Wilson, William Julius. 2009. *More Than Just Race: Being Black and Poor in the Inner City (Issues of Our Time)*. New York: Norton.

Yeung, Wei-Jun Jean and Kathryn M. Pfeiffer. 2009. "The Black–White Test Score Gap and Early Home Environment." *Social Science Research* 38: 412–37.

Author Biography

David M. Quinn is a doctoral student at the Harvard Graduate School of Education. His research interests relate to measuring, explaining, and ending educational inequity; he is particularly interested in how teachers and teaching practice can improve student learning and close achievement gaps.