American Educational Research Journal

http://aerj.aera.net

The Hispanic-White Achievement Gap in Math and Reading in the Elementary Grades

Sean F. Reardon and Claudia Galindo

Am Educ Res J 2009 46: 853 originally published online 10 April 2009

DOI: 10.3102/0002831209333184

The online version of this article can be found at:

http://aer.sagepub.com/content/46/3/853

Published on behalf of

American Educational Research Association

and

SAGE

http://www.sagepublications.com

Additional services and information for *American Educational Research Journal* can be found at:

Email Alerts: http://aerj.aera.net/alerts

Subscriptions: http://aerj.aera.net/subscriptions

Reprints: http://www.aera.net/reprints

Permissions: http://www.aera.net/permissions

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

American Educational Research Journal
September 2009, Vol. 46, No. 3, pp. 853–891
DOI: 10.3102/0002831209333184
© 2009 AERA. http://aerj.aera.net

The Hispanic-White Achievement Gap in Math and Reading in the Elementary Grades

Sean F. Reardon
Stanford University
Claudia Galindo
University of Maryland, Baltimore County

This article describes the developmental patterns of Hispanic-White math and reading achievement gaps in elementary school, paying attention to variation in these patterns among Hispanic subgroups. Compared to non-Hispanic White students, Hispanic students enter kindergarten with much lower average math and reading skills. The gaps narrow by roughly a third in the first 2 years of schooling but remain relatively stable for the next 4 years. The development of achievement gaps varies considerably among Hispanic subgroups. Students with Mexican and Central American origins—particularly first- and second-generation immigrants—and those from homes where English is not spoken have the lowest math and reading skill levels at kindergarten entry but show the greatest achievement gains in the early years of schooling.

Keywords: achievement gap, Hispanic education, elementary schools

Despite the rapid growth and the substantial diversity of the Hispanic school-age population in the United States, we have relatively little detailed and systematic knowledge regarding achievement patterns among

---

Sean F. Reardon is an associate professor of education and (by courtesy) sociology at the Stanford University School of Education, CERAS Building, 520 Galvez Mall, #526, Stanford, CA 94305-3084; e-mail: *sean.reardon@stanford.edu*. His research focuses on the causes, patterns, and consequences of social and educational inequality.

Claudia Galindo is an assistant professor in the Language, Literacy, and Culture Program, University of Maryland, Baltimore County, 1000 Hilltop Circle, Academic IV (A Wing), Room 403, Baltimore, MD 21250; e-mail: *galindo@umbc.edu*. Her research interests focus on educational inequality, particularly Latino students' school experiences, as well as their educational and social emotional outcomes. Her main fields of study include sociology of education, educational policy, and immigration.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

various Hispanic subgroups—particularly in comparison to the extent of available information on achievement patterns among Black students. This is not to say that we have no evidence regarding the achievement of Hispanic students. Recent cross-sectional national studies clearly indicate that the educational outcomes of Hispanic students in U.S. schools lag, on average, well behind those of non-Hispanic White and Asian students and, in some cases, behind those of non-Hispanic[1] Black students (Kao & Thompson, 2003; Kohler & Lazarín, 2007; Lee & Burkham, 2002; Schneider, Martinez, & Owens, 2006). Compared to White and Black children, Hispanic children have lower levels of school readiness at the start of kindergarten (Duncan & Magnuson, 2005; Fryer & Levitt, 2004; Reardon, 2003; Rumberger & Arellano, 2004; Zill, Collins, West, & Hausken, 1995). High school completion rates for Hispanic students are substantially lower than those for either White or Black students (Kaufman, Alt, & Chapman, 2001; Padron, Waxman, & Rivera, 2002; Perreira, Harris, & Lee, 2006). Likewise, Hispanic students are less likely than White students to attend and graduate from college (Cameron & Heckman, 2001; Stoops, 2004) and are more likely to be enrolled in 2-year colleges than in 4-year colleges (Fry, 2004).

Nonetheless, most studies of achievement patterns treat Hispanic students as a single, undifferentiated category, typically comparing their averages as a whole with those of White students. Until recently, nationally representative studies comprising sizable samples of Hispanic students have included data only on 4th graders or students in 8th–12th grade,[2] thereby limiting our knowledge about the development of Hispanic-White achievement gaps during elementary school.

Given the lack of existing evidence describing national patterns of Hispanic students' achievement in the elementary grades, we provide here a detailed descriptive analysis of the development of Hispanic-White math and reading achievement gaps in elementary school, paying attention to variation in these patterns among Hispanic subgroups. We use K–5 test score data from a nationally representative sample of students who were in kindergarten the fall of 1998. Although a full accounting of the causes of the observed patterns is certainly necessary, we do not aim in this article to explain the causes of these gaps nor to suggest or evaluate remedies. Just as in medicine, where epidemiological documentation may stimulate the discovery of a cure, so too in educational research, a detailed description of the development of achievement gaps may lead to a better understanding of their causes and solutions.

The article proceeds as follows. We begin with a brief review of existing evidence regarding trends and patterns in Hispanic-White achievement gaps in elementary grades, drawing primarily on studies that use nationally representative samples. In the second section, we describe the data and measures on which we rely for the results. We include here a discussion of the methodological and measurement issues involved in assessing the magnitude and development of achievement gaps, because these issues

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

underlie all subsequent results and discussion. In the third section, we describe patterns in the development of Hispanic-White math and reading disparities through elementary school; here, we report gaps for Hispanic subgroups, as defined by national origin and immigrant generational status.[3] We also investigate the extent to which these gaps differ among Hispanic subpopulations, as defined by language use and socioeconomic status. We conclude with a discussion of the implications of our findings.

Educational Achievement Patterns of Hispanic Students

The best recent data on the size of Hispanic-White achievement gaps in the elementary grades comes from the National Assessment of Educational Progress (NAEP). The so-called main NAEP assessments have been given to statewide and nationally representative samples of students in Grades 4, 8, and 12, roughly every 2 years since the early 1990s. Using publicly available NAEP data, we estimate that the Hispanic-White achievement gap in 2007 was roughly three quarters of a standard deviation in math and reading in 4th and 8th grade. These gaps have declined slightly in the last decade, by roughly one-tenth of a standard deviation.[4]

Although NAEP data are useful for examining trends, they do not provide a detailed description of how gaps develop as students progress through school—because such data are available for only a few grade levels and because the data are based on repeated cross-sections of grade cohorts (as opposed to longitudinal data on cohorts). NAEP data suggest that Hispanic-White gaps do not substantially change in the late elementary and middle school years (i.e., the 4th- and 8th-grade gaps are similar in magnitude); although changes in the composition of the cross-sectional grade cohorts potentially complicate such comparisons. Studies with repeated measures of achievement for the same sample of students provide the best information on how achievement gaps develop.

The best evidence on the development of Hispanic-White achievement gaps comes from three studies containing longitudinal test score data from large representative samples of Hispanic and White elementary school students.[5] First is the Early Childhood Longitudinal Study–Kindergarten Class of 1998–1999 (ECLS-K), a study of a nationally representative sample of students who were in kindergarten in the fall of 1998 (National Center for Education Statistics, 2001). The ECLS-K sample was tested in math and reading at multiple times from 1998 to 2004. Fryer and Levitt (2006) report estimated Hispanic-White gaps from kindergarten through third grade using these data. Second is a study based on data from five cohorts of students in North Carolina who were in Grades 3–8 from 1995 to 2004 (Clotfelter, Ladd, & Vigdor, 2006). The last study uses data from two cohorts of the 1991–1994 Prospects study, one followed from first to third grade and another from third to sixth grade (Borman, Stringfield, & Rachuba, 2000). Each of these studies provides repeated estimates of Hispanic-White achievement gaps in math and reading for cohorts of students in elementary school in the last two decades.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

With the exception of the first-grade cohort from the Prospects study, each study generally finds that Hispanic-White gaps in math and reading narrow somewhat as children progress through elementary school.[6] The ECLS-K data show that the most rapid narrowing occurs in the first 2 years of schooling (Fryer & Levitt, 2006), although the North Carolina and Prospects data indicate that the gaps continue to narrow through eighth grade (Borman et al., 2000; Clotfelter et al., 2006). While there are differences in the sizes of the estimated gaps across studies, these may result from differences in the cohorts of students included and differences in the tests used in each study.

Diversity of the Hispanic Student Population

Although the aforementioned studies provide some evidence that Hispanic-White achievement gaps narrow slightly during elementary school, each study treats Hispanic students as a homogeneous group. Given the size of the Hispanic student population—one sixth of the school-age population and more than one fifth of public elementary school enrollments (National Center for Education Statistics, 2002a)—a fuller account of Hispanic students' achievement patterns would consider their substantial diversity.

The Hispanic student population is diverse along several key dimensions, including national origin, family immigration history (including immigrant generation, context of reception, and length of time in the United States), socioeconomic status, and linguistic and cultural characteristics. With regard to national origin, Mexicans are by far the most predominant Hispanic group in the United States, representing 59% of the Hispanic population. Next in size are Puerto Ricans (10%), Central Americans (including Dominicans; 7%), South Americans (4%), and Cubans (3.5%) (Guzman, 2001; Ramirez, 2004). With regard to immigration history, the Hispanic population's diversity is reflected in its heterogeneity with regard to nativity and immigrant generational status. Roughly 60% of the total U.S. Hispanic population in 2000 was native-born (Ramirez, 2004). Among Hispanic school-age children, 84% are native-born, although most of them (53% of all Hispanic school-age children) have at least one parent born outside the United States or in Puerto Rico. Less than one third of Hispanic children are third-generation immigrants (or higher), meaning that both parents were born in the United States (Hernandez, 2006; Planty et al., 2008).

With regard to socioeconomic status, the Hispanic population is, in general, economically disadvantaged in comparison to the non-Hispanic U.S. population: Hispanic median family income in 2003 was $33,000, only 69% of White median family income (DeNavas-Walt, Proctor, & Mills, 2004). Likewise, 27% of Hispanic school-age children lived below the poverty line in 2000, compared to 10% and 33% of White children and Black children, respectively (Planty et al., 2008). Moreover, economic conditions vary substantially across Hispanics of different national origins. Child poverty rates

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

are particularly high among Dominicans (35%), Mexicans (28%; 36% among first-generation Mexicans), and Puerto Ricans (33%) but are much lower for Cubans (15%) and South Americans (17%) (Lichter, Qian, & Crowley, 2005).

The diversity of the Hispanic student population is also evident in their varying levels of proficiency and use of English and Spanish. Nationally, 21% of Hispanics (and 31% of school-age Hispanic children) report that they speak only English at home, whereas 41% of Hispanics (and 18% of school-age Hispanic children) report that they do not use English at home and do not speak English well (Planty et al., 2008; Ramirez, 2004). Language use and proficiency, however, vary considerably among Hispanic subgroups. Among Hispanics of Central American and Dominican origin, only 43% and 46% are proficient in English, compared to 73% of Puerto Ricans, for example (Hakimzadeh & Cohn, 2007; Ramirez, 2004).

The substantial heterogeneity of the Hispanic student population suggests that Hispanic achievement patterns are not well characterized by a single measure of the Hispanic-White gaps. As such, in this article we describe the developmental patterns of achievement gaps for various Hispanic subgroups, to provide a more nuanced description of patterns of Hispanic achievement.

Explanations of Hispanic-White Achievement Gaps

Although our primary purpose in this article is descriptive, we find it useful to consider the possible explanations for Hispanic-White achievement disparities. In general, research suggests that such differences are attributable to three major causes—family background and socioeconomic status, English proficiency, and school quality. Of these, the strongest determinant of achievement is family socioeconomic status (Rothstein, 2004), which affects children's educational opportunities, access to resources, role models, and verbal interactions (Gándara, Rumberger, Maxwell-Jolly, & Callahan, 2003; Hao & Bonstead-Bruns, 1998; Schmid, 2001). Because Hispanic children are three times more likely than White children to grow up in poverty (Lichter et al., 2005), four times more likely to have parents who did not complete high school (Galindo & Reardon, 2006; Smith, 2006), and one third more likely to have moved within the past year (Neut, 2006), they have, on average, fewer educational resources and opportunities than White children do. Indeed, Fryer and Levitt (2006) find that such family socioeconomic factors largely—though not entirely—account for Hispanic-White achievement gaps at the start of kindergarten.

A second major factor that may affect Hispanic-White achievement gaps is students' linguistic environment and language proficiency (Gándara et al., 2003; Padilla & Gonzalez, 2001; Winsler, Diaz, Espinosa, & Rodriguez, 1999). English-language learners and limited-English proficiency students have significantly lower achievement test scores than do native English-speaking students (Gándara et al., 2003; Portes & Schauffler, 1996; Thomas

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

& Collier, 2002). The effect of students' English proficiency on achievement may depend, however, on school characteristics. In schools where English is the only language of instruction, Hispanic students need at least minimum English skills to understand instructional content, participate in meaningful learning interactions, and engage in inquiry processes that further learning (Henderson & Landesman, 1992; Rosenthal, Baker, & Ginsburg, 1983). With higher levels of oral English proficiency, English-language learners have better academic language and a wider repertoire of language-learning strategies, and they are able to better participate in meaningful learning interactions (Saunders & O'Brian, 2007). As a result, language proficiency patterns may interact with school curricula and instructional practices to affect the development of achievement gaps. Parental English proficiency may also affect student learning. Hispanic parents—particularly those who do not speak fluent English and who did not attend school in the United States—may be less familiar and less comfortable with the practices, expectations, and institutions of U.S. schooling than are native-born parents and English-speaking parents. These differences may therefore lead to differences in student learning because the parents of Hispanic students may be less likely to monitor, support, and intervene in their students' schooling experiences.

Third, Hispanic and White students' school environments differ, on average, in ways that may affect learning. Hispanic students are more segregated from White students than Black students are (Orfield & Yun, 1999), and they are likely to attend schools with few experienced teachers (Galindo & Reardon, 2006) and with high concentrations of poor and non-English-proficient students (Crawford, 1997; Schmid, 2001; Van Hook & Balistreri, 2002). To the extent that the structural and compositional features of schools (such as poverty and racial/ethnic composition) are related to access to resources, quality teachers, and academic culture, Hispanic students may be particularly disadvantaged.

Data and Method

Data

The data we use come from the ECLS-K, sponsored by the National Center for Education Statistics. The ECLS-K contains data on a nationally representative sample of approximately 21,400 students from the kindergarten class of 1998–1999 (thus, representing a cohort born in 1992–1993). Students in the sample were assessed in reading and mathematics skills at six points from 1998 to 2004 (fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, and spring 2004). In addition to these cognitive developmental measures, the ECLS-K data include information gathered from parents, teachers, and school administrators regarding family, school, community, and student characteristics.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

To disaggregate the Hispanic student population, we categorize Hispanic students using four variables: national/regional origin, immigrant generational status, socioeconomic status, and language used at home. We describe the construction of each variable below. Although the measures of immigrant generational status, national/regional origin, and language used at home that we construct from information available in the ECLS-K data are not ideal, they are certainly better than previously available data.

National/regional origin. Student classification is based on parent survey responses. Students were classified as having national origins in Mexico, Puerto Rico, Cuba, South America, Central America (including the Dominican Republic), and elsewhere. This last category (*other Hispanic origin*) includes a small and heterogeneous group of students with ancestries in Spain, Brazil, Guyana, and Dominica, for example. This category also includes students for whom country of birth information is missing and whose parents defined them as members of an *other Spanish/Hispanic/Latino group* in the question about Hispanic group membership.[7]

Immigrant generational status. We categorize ECLS-K Hispanic students as first-, second-, or third-(plus)-generation on the basis of a set of questions that indicate where they and their parents were born (i.e., from the kindergarten, first-, and third-grade parent surveys). We classify as first-generation students those students born outside of the U.S. whose mothers (or fathers, for those whose mothers' birthplace was not reported) were born outside of the U.S.; island-born Puerto-Rican students are also classified as first-generation students. Second-generation students are those born in the U.S. whose mothers (or fathers, for those whose mothers' birthplace was not reported) were born outside of the U.S. Finally, we classify as third-generation students those born to a U.S.-born parent, regardless of where the student was born (note that when we refer to *third-generation students*, we are including those who are beyond third generation).[8] We report achievement gaps by generational status for Mexican-origin students only, because sample sizes for other groups are too small.

Language used at home. Because the ECLS-K does not include direct measures of English proficiency for all students, we use as a proxy a measure of the language (or languages) spoken by the student and his or her parents at home in the fall of kindergarten. Parents were asked how often each parent speaks Spanish to the child and how often the child speaks Spanish to each parent (*never* to *very often*). We averaged the parent responses to these four questions ($\alpha = .96$) and then rounded this average to the nearest whole number, yielding a categorical variable describing the primary language used at home: only English, primarily English, primarily Spanish, or only Spanish.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Socioeconomic status. For the ECLS-K, a continuous measure of socioeconomic status was created; it was based on a composite of the educational attainment of the student's mother and father, the occupation of the student's mother and father, and family income (National Center for Education Statistics, 2002b). We averaged the kindergarten and first-grade composite measures for each student and classified students by quintiles of this measure.

The ECLS-K Math and Reading Assessments

The ECLS-K direct cognitive assessments are individually-administered, oral, untimed, adaptive tests of math and reading skills. The content areas of the tests are based on the NAEP fourth-grade content areas, adapted to be age appropriate at each assessment. The assessments were administered by trained ECLS-K assessors, and were scored using a three-parameter Item Response Theory (IRT) model. Details of the assessments are provided in the ECLS-K psychometric reports (Pollack, Narajian, Rock, Atkins-Burnett, & Hausken, 2005; Pollack, Rock, et al., 2005; Rock & Pollack, 2002).

Students who were proficient in oral English at the time of assessment were administered the ECLS-K math and reading assessments in English. Students who were not proficient in oral English but proficient in oral Spanish were administered the ECLS-K math assessment in Spanish but were not administered the reading assessment.[9] This has implications for our ability to estimate trends in math and reading test scores. In the early waves of the ECLS-K data collection, many Hispanic students were not fluent enough in oral English to be assessed in reading (29% of all Hispanic students in the study, including 42% of Mexican-origin students and 77% of first-generation Mexican-origin students). The proportion of oral English-proficient Hispanic students grows over time—to 80% by the spring of kindergarten, 90% by the spring of first grade, and 99% by the spring of third grade. Furthermore, students not proficient in English have lower average reading skills in English than do students proficient in oral English. As a result, trends in the mean reading scores of those Hispanic students with reading scores are confounded by changes in the population of Hispanic students who are represented in the sample of students with test scores. Because non-English-proficient Hispanic students could take the math test in Spanish, there is no such bias in the mean math score estimates (except for Asian students, 22% of whom were not proficient in English at Wave 1 and for whom no home-language versions of the tests were available).[10]

To avoid these confounding patterns, we focus here on estimating reading achievement trends for only the subpopulation of students who were proficient in spoken English at the start of kindergarten. Thus, the Hispanic-White reading gaps reported here almost certainly understate the magnitude of the true Hispanic-White reading gaps, because they are based on only the 71% of Hispanic students proficient in oral English in kindergarten. Moreover, the Hispanic students not proficient in English in the fall of Kindergarten

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

differ substantially on key demographic dimensions from those who were proficient: Nonproficient students were disproportionately Mexican (79% compared to 64% of proficient students) and Central American (13% compared to 8%), disproportionately first- and second-generation immigrants (11% and 83% compared to 1% and 48%, respectively), disproportionately from homes where English was not the primary language (90% compared to 31%), and disproportionately poor (76% were in the lowest socioeconomic quintile, compared to 31% of English-proficient students). We caution readers to keep this in mind throughout this article. As a partial assessment of the extent to which this restriction may affect our conclusions, we include supplemental analyses examining the differences in skills between those students in our analytic reading sample and those excluded because they were not proficient in oral English at the start of schooling.

Sample and Descriptive Statistics

We use a subsample of the full ECLS-K sample for our analyses. Our analytic subsample excludes immigrant White and Black students (i.e., first- and second-generation immigrants), to ensure that our comparisons focus on the differences between Hispanics and non-immigrant White and Black students. Asian students—the majority of whom are first- and second-generation immigrants—are included regardless of immigrant generation.

Second, to ensure that our trend data are not affected by attrition from the ECLS-K sample, we restrict our analyses of achievement gaps to a subsample of the ECLS-K students who were present in the ECLS-K sample at Waves 1, 2, 4, 5, and 6.[11] By design, the ECLS-K study did not follow all students through Wave 6; when students changed schools, only a random subsample of those changing schools were followed. We use the ECLS-K longitudinal sampling weight C1_6FC0, a child-specific weight designed for analyses using the sample of students present at Waves 1, 2, 4, 5, and 6 (such as our analyses here; Tourangeau et al., 2006), to produce unbiased estimates for the population of U.S. kindergarten students in the fall of 1998.[12]

Finally, we restrict the math and reading analyses to students who have valid math and reading scores in Wave 1 (fall of kindergarten). In the case of the reading assessment, this restriction limits our analyses to students who were proficient in oral English in the fall of kindergarten (71% of all Hispanic students). Table 1 reports the total number of students, by subgroup, used in the math and reading gap analyses. The sample of students that we use in the math analyses is representative of the Hispanic school-age population of the United States.[13] We do not report gap estimates for subgroups comprising fewer than 40 students.

Measuring Achievement Gaps

To compare the magnitude of achievement gaps across grades and ages, the gaps must be measured in a comparable metric at each time point. In the

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Table 1

Analytic Sample Sizes of the Early Childhood Longitudinal Study–Kindergarten Class, by Race, Hispanic National Origin, and Mexican Immigrant Generational Status

	Analytic Sample Size	
Race / National Origin	Math	Reading
White, not Hispanic, third generation	5,119	5,115
Hispanic, any race	1,871	1,303
Mexican origin	1,045	612
Mexican, first generation	77	14
Mexican, second generation	662	308
Mexican, third generation	295	284
Mexican, unknown generation	11	6
Cuban origin	58	47
Puerto Rican origin	119	109
Central American origin	174	100
South American origin	87	73
Other Hispanic origin	101	97
National origin unknown	287	265
Black, not Hispanic, third generation	894	894
Asian	489	489
Other race	538	538
Total	8,911	8,339

ECLS-K study, the cognitive skill of each student i in subject s (math or reading) at each assessment wave t, denoted θ_{ist}, was estimated using a model based on item response theory (Lord & Novick, 1968; Pollack, Narajian, et al., 2005). The θ scores produced by the model are measured in a common metric (within subjects) across assessment waves (i.e., scores are not normed within grades) meaning that, for example, a first-grade student and a fifth-grade student with the same θ have the same skill level. These θ scores are the basis for our estimates of Hispanic-White achievement gaps.[14]

In general, there are three ways to measure achievement gaps—by differences in mean scores, by differences in standardized scores, and by so-called metric-free gap measures (Reardon & Robinson, 2007). We begin by reporting Hispanic-White gaps in different metrics, noting the interpretation of each metric. Throughout the remainder of the article, however, we report gaps using only standardized score differences, which allows for simple interpretation and comparability with other research. The measures that we use are described below; estimation details are in the appendix.

Mean score differences. The most obvious way of measuring an achievement gap between two groups is to define the gap as the difference in their mean test scores. If scores are measured in the same metric across grades,

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

then we can, in principle, compare the difference in mean scores at one grade or age with the mean difference at another. Comparing the magnitude of achievement gaps in this way, however, requires a test that measures cognitive skill in an interval-scaled metric, so that a difference of one point in mean scores between two groups has the same meaning as a difference of one point in mean scores between two other groups or between the same two groups at another time (i.e., it corresponds to the same-size gap in cognitive skills), regardless of where the group means lie on the test score metric.[15] A well-constructed test can generally be assumed to be ordinal scaled (i.e., higher scores correspond to higher levels of cognitive skill), but the assumption of interval scaling is harder to justify (Phillips, 2000; Reardon, 2008). If a test is not interval scaled, however, the magnitude and meaning of a difference in groups' mean scores will depend on where on the scale the difference lies, thus rendering inferences about achievement gaps highly dependent on the test metric (see, e.g., Reardon, 2008; Selzer, Frank, & Bryk, 1994). Although some of the literature on achievement gaps has relied on unstandardized mean differences in test scores (Hanushek & Rivkin, 2006; LoGerfo, Nichols, & Reardon, 2006; Murnane, Willett, Bub, & McCartney, 2006), Reardon (2008) illustrates that analyses based on mean score differences may be sensitive to the test metric used.

Standardized score differences. Because of this dependence—and because it is not obvious that the θ scores or any test scores are measured in an interval-scaled metric (see Ballou, 2008; Phillips, 2000; Reardon, 2008)—we also compute standardized score differences, dividing the θ scores by their pooled standard deviation at each wave and estimating the Hispanic-White gap at each wave in standard deviation units. These standardized score differences are analogous to effect sizes. They have the advantage of being less sensitive (than mean score differences) to the metric used to measure the gaps; in addition, standardized gaps and their standard errors are easily computed from published summary data and so allow rough comparability across different tests. Moreover, standardized gap measures are widely used in the literature (see, e.g., Clotfelter et al., 2006; Fryer & Levitt, 2004, 2006; Grissmer, Flanagan, & Williamson, 1998; Hedges & Nowell, 1999; Neal, 2006; Phillips, Brooks-Gunn, Duncan, Klebanov, & Crane, 1998; Reardon & Galindo, 2006) and so have the advantage of comparability with prior work.

Standardized gap measures have three disadvantages, however. First, they do not make clear how test score differences correspond to interpretable differences in math and reading skills. Second, standardized gap measures confound changes in the difference in the means of two groups with changes in the variation in test scores within groups. They are therefore measures of relative achievement differences—that is, they measure the size of the gap relative to the amount of variation in test scores within each group. If we are interested in absolute mean difference, then standardizing may obscure the

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

information that we desire. Third, measurement error in test scores will tend to inflate the variance of the test score distributions, meaning that the achievement gaps measured in standard deviation units will be biased toward zero.[16] If the gaps at different grades, ages, or cohorts are measured with tests that have different amounts of measurement error, then the amount of bias will not be the same in each measure of the gap, thus leading to potentially erroneous inferences regarding patterns or trends in the magnitudes of the gaps. If the reliability of the test is known, however, then estimated gaps can be corrected for measurement error.

Metric-free gap measures. Standardized gap measures may be sensitive to the distribution of the test scores—which may be affected by the test metric used, because a nonlinear transformation of the test metric will alter the distributions of each group's test scores differently. So-called metric-free measures of achievement gaps, however, rely only on the ordinality of the test metric, thus requiring no assumption of interval scaling (Ho & Haertel, 2006). One such measure, for example, is the probability that a randomly chosen Hispanic student has a score higher than that of a randomly chosen White student (denoted $P_{h>w}$); the less overlap there is of the White and Hispanic test score distributions, the lower $P_{h>w}$ will be. A second metric-free measure is what Ho and Haertel refer to as the pseudo-effect size; this is the effect size (or standardized gap measure) that would correspond to $P_{h>w}$ if both groups had normal test score distributions (Ho & Haertel, 2006).

Because the metric-free measures are essentially measures of the overlap between two distributions, they share the disadvantages of the standardized measures. They do not make clear the concrete differences in skills between two groups; they are measures of relative rather than absolute difference; and they are biased toward zero by measurement error. Unlike the standardized measures, however, which are somewhat sensitive to violations of the interval-scaling assumption, metric-free measures are invariant under any monotonic transformation of the test metric, thus obviating the need for assumptions about the metric's interval nature. Despite this advantage over standardized gap measures, metric-free measures have not yet been widely used (for recent examples of their use, see Ho & Haertel, 2006; Neal, 2006; Reardon, 2008).

Estimating Achievement Gap Trends

We estimate achievement gaps between Hispanic and White students at each wave of the ECLS-K study. In addition to estimating the gaps at each wave, we fit piecewise trend lines through the standardized gaps of Waves 1–6 to provide a simpler visual summary of their developmental trends and to test hypotheses regarding differences in the sizes of the gaps over time and among Hispanic subgroups. These trends are fitted with a change in slope at the end of first grade; the two slopes indicate the average linear

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

trend in the achievement gap during kindergarten and first grade and the average linear trend in the achievement gap from the end of first grade through fifth grade. The appendix provides details on the estimation of wave-specific achievement gap measures and piecewise trend lines.

Results

Measures of Hispanic-White Gaps in Kindergarten Through Fifth Grade

Table 2 shows the estimated Hispanic-White gaps in math and reading as measured using each metric described above (mean θ score differences, standardized score differences, metric-free probabilities, and metric-free pseudo-effect sizes). In addition, because the gap measures are biased toward zero by measurement error in the ECLS-K tests (except for the mean score differences), we report gaps adjusted for measurement error, under assumed test reliabilities of .70, .80, .90, and 1.00.[17] We report standard errors for the mean score differences and standardized differences to provide some sense of the precision regarding the difference estimates.[18] Because our sample size is much smaller in Wave 3 (fall first grade), owing to the ECLS-K design, we report only the estimates from the five waves where we have test scores for our full analytic subsample (gaps estimated from the Wave 3 subsample are similar in magnitude to those at Waves 2 and 4 but are relatively uninformative because of their much larger standard errors).

In the fall of kindergarten, there are large math and reading gaps between Hispanic and White students. Although the magnitude of the mean θ difference is difficult to interpret concretely, the other metrics provide interpretable gap measures. The estimated math gap is between 0.77 and 0.92 standard deviations (depending on the amount of assumed measurement error). The reading gap is about one third smaller (but recall that the reading gap is estimated for only the 71% of students in the sample who were proficient in oral English in fall Kindergarten).[19] The "metric-free" probability measure yields probabilities of 35% and 29% that a randomly chosen Hispanic student has a higher math score and reading score, respectively, than does a randomly chosen White student in the fall of kindergarten. Finally, note that the pseudo-effect size measures are similar to the standardized differences, reflecting the fact that the θ distributions are approximately normal.

The Hispanic-White math gap narrows from kindergarten to fifth grade, regardless of the metric used to measure it, although the timing of the narrowing depends somewhat on the metric. In each metric, the gap narrows sharply in kindergarten and first grade but then flattens. In the mean θ difference metric, the gap remains flat through fifth grade, whereas it narrows from third to fifth grade in the other metrics, a reflection of the fact that the standard deviations of the within-group test score distributions increase from third to fifth grade, thus making the relative gap smaller.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Table 2

Hispanic-White Math and Reading Test Score Gaps, Kindergarten Through Fifth Grade, by Gap Measure and Wave

	Math					Reading				
	Fall K	Spring K	Spring First	Spring Third	Spring Fifth	Fall K	Spring K	Spring First	Spring Third	Spring Fifth
Mean θ score differences	−.36	−.31	−.23	−.22	−.22	−.25	−.16	−.13	−.12	−.13
	(.03)	(.03)	(.02)	(.02)	(.02)	(.03)	(.03)	(.03)	(.02)	(.02)
Standardized differences										
$r = .7$	−.92	−.80	−.67	−.68	−.60	−.61	−.41	−.35	−.43	−.45
	(.07)	(.07)	(.06)	(.06)	(.06)	(.07)	(.07)	(.07)	(.07)	(.07)
$r = .8$	−.86	−.75	−.63	−.64	−.56	−.57	−.38	−.32	−.40	−.42
	(.07)	(.07)	(.06)	(.06)	(.06)	(.07)	(.07)	(.07)	(.07)	(.07)
$r = .9$	−.81	−.71	−.59	−.60	−.53	−.54	−.36	−.31	−.38	−.40
	(.07)	(.07)	(.06)	(.06)	(.06)	(.07)	(.07)	(.07)	(.07)	(.07)
$r = 1.0$	−.77	−.67	−.56	−.57	−.50	−.51	−.34	−.29	−.36	−.38
	(.06)	(.06)	(.05)	(.05)	(.05)	(.06)	(.06)	(.06)	(.06)	(.06)
Metric-free probabilities										
$r = .7$	.26	.28	.31	.31	.33	.32	.38	.39	.37	.36
$r = .8$	.27	.30	.32	.32	.34	.33	.39	.40	.38	.37
$r = .9$	.28	.31	.33	.33	.35	.34	.40	.41	.39	.38
$r = 1.0$	.29	.32	.34	.34	.36	.35	.41	.42	.40	.39
Metric-free pseudo-effect sizes										
$r = .7$	−.92	−.81	−.72	−.69	−.61	−.66	−.42	−.39	−.48	−.52
$r = .8$	−.86	−.75	−.67	−.64	−.57	−.62	−.39	−.36	−.43	−.47
$r = .9$	−.81	−.71	−.63	−.61	−.54	−.59	−.36	−.32	−.39	−.43
$r = 1.0$	−.77	−.67	−.60	−.57	−.51	−.56	−.34	−.30	−.36	−.39

Note. Standard errors in parentheses.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

In reading, the Hispanic-White gap narrows sharply in kindergarten and slightly in first grade but then remains flat in the mean θ difference metric but widens somewhat in the other metrics. Here the explanation for the divergence of these trends is a result of the changing standard deviation of test scores, but in this case, the standard deviation narrows from first to third grade. The narrowing standard deviation of the θ scores means that a constant difference in θ corresponds to a widening standardized difference.

The disparities in the developing patterns of the Hispanic-White gaps among different measures are not dramatic here, although they do indicate the potential sensitivity of our conclusions to the choice of a test metric. For the reasons outlined above, there is no "best" metric; however, each of those we describe has some merit. In the interest of space and parsimony, we rely on the standardized difference gap measures throughout the remainder of this article. These have the advantage of being familiar and roughly comparable across studies, and they yield descriptions of the development of gaps that are nearly identical to those based on metric-free measures. For simplicity, rather than report estimates under a range of assumed test reliabilities, we report estimates based on the assumption that the test scores contain no measurement error. These estimates can be divided by the square root of any assumed reliability to recover unbiased gap estimates. Moreover, as long as the reliability of the tests is constant across the waves of the test administration, the development patterns of the gaps are not biased by ignoring measurement error, because all gap estimates will be biased by measurement error in the same direction and by the same proportional amount.

Comparing Math and Reading Gaps

In Table 2, we note that the Hispanic-White reading gap is roughly one third smaller than the math gap, but these gaps are not strictly comparable, because the reading estimates are based only on students who were English proficient in the fall of Kindergarten. In Table 3, we report estimated math and reading gaps for two groups of students—those who were and were not deemed English proficient in the fall of Kindergarten (based on whether they took the math test in English or Spanish)—to facilitate a more appropriate comparison of the math and reading gaps. For math, we estimate these gaps at each wave, but for reading we can estimate only the gaps in Waves 5 and 6 (spring third grade and spring fifth grade) because not all nonproficient students were assessed in reading in the earlier waves.

Table 3 suggests that the Hispanic-White math and reading gaps in the full population are similar in magnitude, at least in third and fifth grade, where they range between 0.50 and 0.60 standard deviations. Likewise, the gaps for initially proficient and nonproficient students are roughly similar in magnitude in third and fifth grade, although the reading gaps are somewhat larger than the math gaps among nonproficient students (by 0.15 to 0.20 standard deviations). Thus, it would be a mistake to conclude from

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Table 3

Estimated Standardized Math and Reading Hispanic-White Test Score Gaps, by Wave and Hispanic Students' Fall Kindergarten English Proficiency

	Math						Reading	
	Fall K	Spring K	Fall First	Spring First	Spring Third	Spring Fifth	Spring Third	Spring Fifth
Proficient	–0.518	–0.450	–0.350	–0.403	–0.403	–0.349	–0.359	–0.372
	(0.060)	(0.060)	(0.102)	(0.054)	(0.057)	(0.056)	(0.064)	(0.062)
Not proficient	–1.394	–1.239	–1.143	–0.964	–0.991	–0.884	–1.140	–1.088
	(0.067)	(0.067)	(0.104)	(0.062)	(0.062)	(0.063)	(0.061)	(0.055)
Difference	–0.876	–0.789	–0.793	–0.561	–0.588	–0.535	–0.781	–0.716
	(0.090)	(0.090)	(0.146)	(0.082)	(0.084)	(0.084)	(0.088)	(0.083)
All Hispanic students	–0.768	–0.675	–0.606	–0.564	–0.571	–0.502	–0.587	–0.580
	(0.057)	(0.056)	(0.094)	(0.048)	(0.049)	(0.049)	(0.055)	(0.053)

Note. Standard errors in parentheses (survey design corrected). Students are classified as *proficient* or *not proficient* on the basis of whether they they were administered the math test in English or Spanish. Gaps are estimated standardized gaps (assuming test reliability = 1.0) between Hispanic students and third-generation non-Hispanic White students. Sample includes all Hispanic and third-generation White students with a valid longitudinal weight (i.e., those described in the math column of Table 1). Estimates are weighted by C1_6FCO.

Table 2 that the Hispanic-White reading gaps are smaller than the math gaps, because Table 3 clearly indicates that the apparent pattern in Table 2 is an artifact of the differences in the samples used for the math and reading estimates (at least at Waves 5 and 6). For the rest of this article, then, we focus on within-subject comparisons over time and across Hispanic subgroups, rather than on comparisons between math and reading.

Standardized Achievement Gaps in Kindergarten Through Fifth Grade

In this section, we describe the estimated math and reading gaps for a variety of subgroups. Specifically, we report estimated gaps using standardized difference measures, by race/ethnic group; for Hispanics, by country/region of national origin; for Mexican-origin students, by immigrant generational status; for Hispanics, by language used in the home; and for Hispanics, by socioeconomic status quintile.

To facilitate presentation of a large amount of data and because our aims are primarily descriptive, we present the estimated gaps in a series of figures rather than tables (exact gap estimates and their standard errors are available upon request). In Figures 1–10, the vertical axis indicates the size of the achievement gap—that is, the difference in standard deviation units between the average score of students in a given group and the average score of the reference group (generally, third-generation White students). For each comparison group (in Figure 1, e.g., Black, Hispanic, Asian, and other), the

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

figure shows six estimates of the achievement gap, corresponding to the six waves of ECLS-K assessments: fall, kindergarten (FK); spring, kindergarten (SK); fall, first grade (F1); spring, first grade (S1); spring, third grade (S3); and spring, fifth grade (S5). For each gap, the vertical error bars indicate the 95% confidence interval around the estimate. Finally, for each group, the figure includes the fitted trend line, indicated by the thick solid line. The fitted trend lines do not correspond exactly to the point estimates of the gaps at each assessment wave; rather, they summarize the general trends in the magnitude of the gaps during the two periods (kindergarten to first grade and first to fifth grade).

Because some of the subgroup sample sizes are rather small, some of our estimates are imprecise (as evident in their large confidence intervals), meaning that some of the estimated trends and patterns may be heavily influenced by sampling variation. Beneath each figure, we report the results of a set of hypothesis tests designed to identify those trends that are most reliable. Specifically, we test two types of hypotheses. First, for each subgroup, we test the null hypothesis that the gap remains constant from kindergarten through fifth grade (i.e., that the slopes of both piecewise trend lines are equal to zero). Second, we test the null hypothesis that magnitude of the gaps and the shape of their trends are the same for each subgroup displayed (e.g., in Figure 1, we test the null hypothesis that the math gap trends are the same for Black, Hispanic, Asian, and other students). In our discussion of the results, we focus on those patterns and trends where we can reject the relevant hypotheses (we use $\alpha = .10$, given the small sample sizes and probability of Type II errors).

Gaps by race/ethnicity. Figures 1 and 2 show kindergarten through fifth grade trends in the differences in average math scores (Figure 1) and reading scores (Figure 2) of Black, Hispanic, Asian, and other students, relative to White students. Most notable here are (a) the steadily increasing Black-White gaps (particularly in math) during the kindergarten through fifth-grade period and (b) the narrowing of the White-Hispanic gaps during kindergarten and first grade, followed by a period of stability from the end of first grade through fifth grade. At the start of kindergarten, Hispanic and Black students have math and reading scores substantially lower than those of White students (but roughly equal to one another). The average Hispanic and Black students begin kindergarten with math scores three quarters of a standard deviation lower than those of White students and with reading scores a half standard deviation lower than those of White students. Six years later, however, Hispanic-White gaps have narrowed by roughly a third, whereas Black-White gaps have widened also by roughly a third. Despite this narrowing, the Hispanic-White gap is a half standard deviation in math, and three eighths in reading, at the end of fifth grade.

The trends in the Hispanic-White gaps are notable for their rapid narrowing in kindergarten and first grade: The estimated math gap declines

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

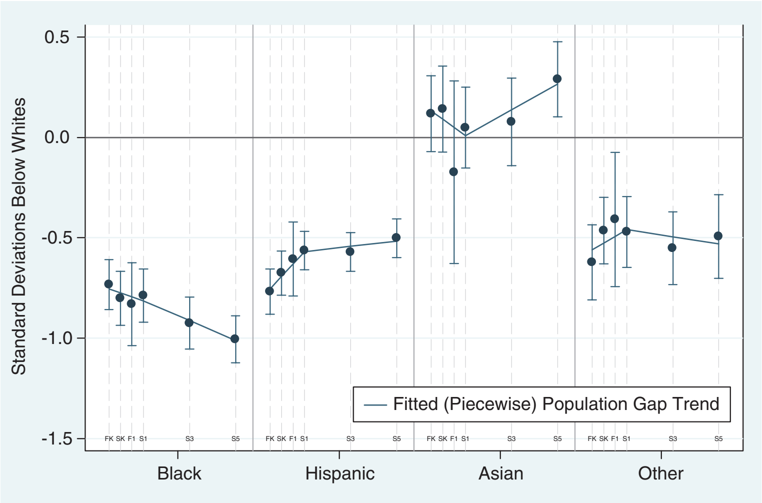

Figure 1. **Trends in estimated math gaps, by race/ethnic group.**
Note. p values for test of null hypothesis that trend for each group is flat: Black, $p = .010$; Hispanic, $p = .009$; Asian, $p = .055$; other, $p = .627$. p value for test of null hypothesis that gap trends are equal for all groups: $p < .001$.

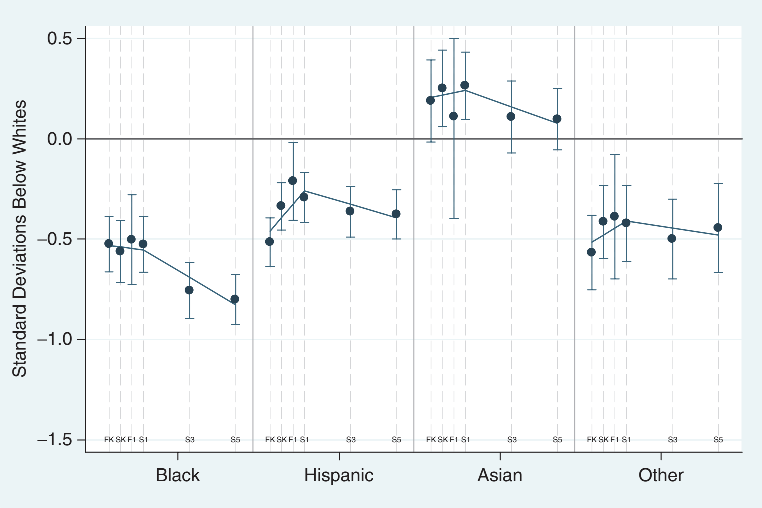

Figure 2. **Trends in estimated reading gaps, by race/ethnic group.**
Note. p values for test of null hypothesis that trend for each group is flat: Black, $p = .022$; Hispanic, $p = .210$; Asian, $p = .017$; other, $p = .507$. p value for test of null hypothesis that gap trends are equal for all groups: $p < .001$.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

from 0.77 to 0.56 standard deviations, and the estimated reading gap from 0.52 to 0.29, in the roughly 18 months between the fall of kindergarten and the spring of first grade. In the 4 years from the spring of first grade through the spring of fifth grade, however, the gaps change very little—narrowing slightly to 0.50 in math and widening slightly to 0.38 in reading. As we will see, this pattern of rapid narrowing at the start of formal schooling, followed by relative stability, is common to most Hispanic subgroups.

Achievement gaps by Hispanic national/regional origin. Figures 3 and 4 describe Hispanic-White achievement gaps disaggregated by Hispanic students' country/region of origin. Several key findings are evident here. First is the considerable heterogeneity among Hispanic national origin groups in the magnitude of achievement disparities, particularly with regard to math achievement. In math, students of Mexican and Central American origins enter kindergarten with achievement scores approximately one standard deviation below those of White students, whereas students of Cuban, Puerto Rican, and South American origins enter kindergarten with scores about a half standard deviation below White students. In reading, the patterns are similar, although the gaps are only half the size of those in math and vary less among national/regional origin groups. As noted above, the reading gaps reported here are based on only the sample of Hispanic students proficient in oral English at the start of kindergarten; so, the gap estimates are likely smaller here than what they would be if we could include all Hispanic students.

A second key finding is that there is some heterogeneity among Hispanic national origin groups in the patterns of achievement gaps from kindergarten to fifth grade. In general, the achievement gaps narrow for most groups in kindergarten and first grade, though not for Puerto Rican students in math or Cuban-origin students in reading. From the spring of first through fifth grade, however, the patterns are more varied. In math, there is little or no change in the size of the achievement gap for students of Mexican and Cuban origin, a gradual narrowing for students of Puerto Rican and South American origin, and a substantial narrowing for students of Central American origin. In reading, the gaps change little for most groups from first through fifth grade.

Achievement gaps by immigrant generational status. To examine differences in achievement patterns by Hispanic immigrant generational status, we focus on Mexican-origin students because they are the only national origin group with sizable samples of first-, second-, and third-generation students. Although we would obtain larger sample sizes within each immigrant generation group if we combined all national origin groups, such an analysis—that is, one that combined students of similar immigrant generations but of different national origins—would potentially confound generational status with the immigration histories and contexts of different national origin groups.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

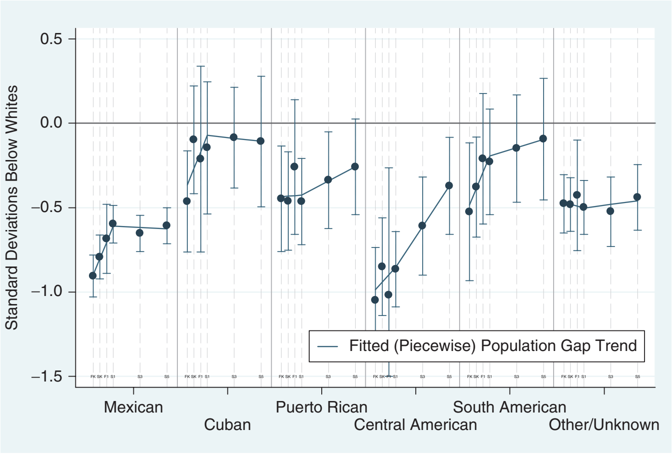

Figure 3. **Trends in estimated Hispanic-White math gaps, by Hispanic national origin.**

Note. p values for test of null hypothesis that trend for each group is flat: Mexican, $p = .006$; Cuban, $p = .297$; Puerto Rican, $p = .258$; Central American, $p = .012$; South American, $p = .043$; other/unknown, $p = .512$; p value for test of null hypothesis that gap trends are equal for all groups: $p < .001$.

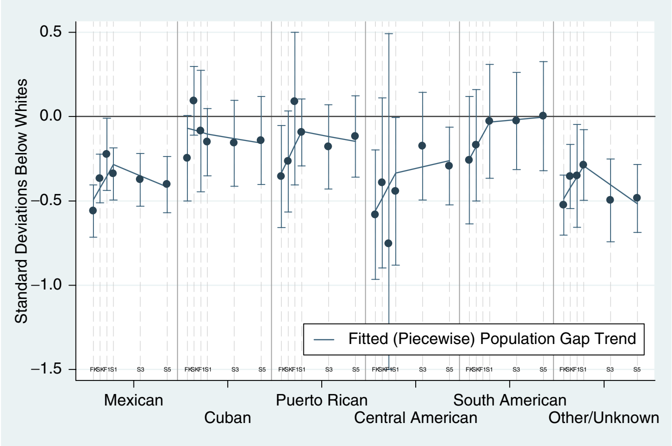

Figure 4. **Trends in estimated Hispanic-White reading gaps, by Hispanic national origin.**

Note. p values for test of null hypothesis that trend for each group is flat: Mexican, $p = .148$; Cuban, $p = .950$; Puerto Rican, $p = .307$; Central American, $p = .826$; South American, $p = .569$; other/unknown, $p = .193$; p value for test of null hypothesis that gap trends are equal for all groups: $p < .001$.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Figures 5 and 6 illustrate the patterns of math and reading scores for second- and third-generation Mexican students (as well as math scores for first-generation Mexican students). In math, first- and second-generation students of Mexican origin enter kindergarten with achievement scores approximately 1.10 standard deviations below those of White students. These gaps are large, much larger, for example, than the Black-White math gaps at the start of kindergarten. Third-generation Mexican students enter kindergarten with math scores 0.46 standard deviations below those of White students. By the spring of first grade, these gaps have narrowed considerably, particularly for first- and second-generation students, who are roughly 0.75 standard deviations below White students at this point. There is relatively little change in the magnitude of the gaps after first grade for any of the groups.

Because our reading gap estimates are based on the sample of students who were proficient in oral English in the fall of kindergarten and because so few first-generation Mexican-origin students were proficient, we estimated reading gaps only for second- and third-generation Mexican-origin students. The patterns for these students in reading are similar to those in math, although the magnitudes of the gaps are smaller. As in math, the gaps are larger for second-generation students than for third-generation students; they narrow sharply in kindergarten and first grade; and they are relatively stable from first to fifth grade.

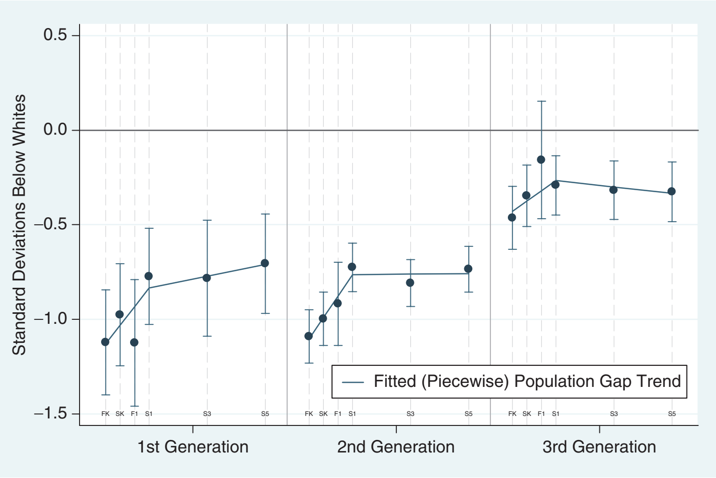

Figure 5. **Trends in estimated Hispanic-White math gaps, by immigrant generational status: Mexicans.**

Note. p values for test of null hypothesis that trend for each group is flat: first-generation Mexican, p = .103; second-generation Mexican, p = .012; third-generation Mexican, p = .228; p value for test of null hypothesis that gap trends are equal for all groups: p < .001.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

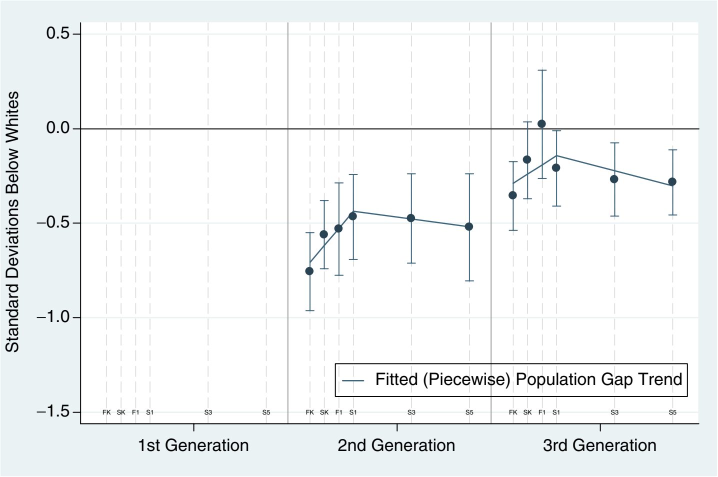

Figure 6. **Trends in estimated Hispanic-White reading gaps, by immigrant generational status: Mexicans.**

Note. p values for test of null hypothesis that trend for each group is flat: second-generation Mexican, $p = .146$; third-generation Mexican, $p = .576$; p value for test of null hypothesis that gap trends are equal for both groups: $p = .001$.

Achievement gaps by language use in the home. Table 4 reports the distribution of Hispanic students' home language use, by country of origin and generational status. Overall, Hispanic students are equally likely to live in English- or Spanish-dominant homes; however, the patterns of home language show sharp differences among Hispanic subgroups. Most students of Mexican and Central American origins, particularly the most recent Mexican immigrants, come from families where Spanish is the predominant or only language spoken in the home. Puerto Rican students and students of Other Hispanic origin, in contrast, come primarily from homes where English is the predominant or only language spoken in the home.

Figures 7 and 8 show Hispanic-White achievement gap trends for Hispanic subgroups defined by language spoken in the students' homes. Two clear patterns are evident: First, students from homes where Spanish is the only language or the predominant language enter kindergarten with lower math and reading skills than those of students from homes where English is the dominant language. Second, the pattern of rapid achievement gains in kindergarten and first grade is most evident for students from homes where Spanish is the only language spoken. This pattern is particularly evident in reading. Nonetheless, despite the rapid gains in math and reading of students from homes where Spanish is the predominant language, these students still score well below those of White students and Hispanic students from English-speaking homes by fifth grade.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Table 4
Home Language Use, by National Origin and Mexican Immigrant Generational Status: Hispanics

	Language Spoken in Home			
Origin / Generational Status	Only English	Predominantly English	Predominantly Spanish	Only Spanish
Mexican origin	24.0	19.2	20.7	36.0
Mexican, first generation	4.7	0.0	15.0	80.4
Mexican, second generation	11.9	13.3	25.1	49.7
Mexican, third generation	51.0	34.2	10.7	4.1
Cuban origin	7.8	26.3	30.6	35.3
Puerto Rican origin	38.5	26.4	20.6	14.5
Central American origin	23.5	11.2	21.2	44.1
South American origin	10.2	28.8	34.0	27.0
Other Hispanic origin	59.7	18.4	8.9	13.0
National origin unknown	61.7	21.0	10.5	6.9
Total Hispanic, any race	31.8	19.8	19.3	29.1

Note. Percentages are weighted by the ECLS-K longitudinal weight C1_6FC0 and include only students in the longitudinal sample who have valid math scores in Wave 1 (i.e., those in the math column of Table 1). Percentages are not reported for the Mexican–unknown generation group, owing to small sample size.

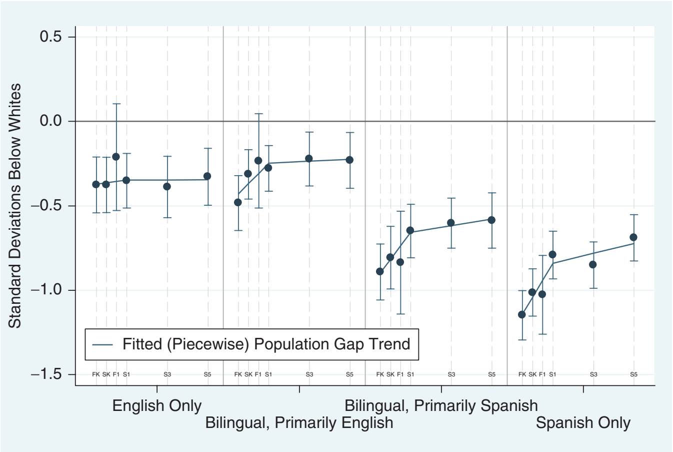

Figure 7. **Trends in estimated Hispanic-White math gaps, by language used in home.**
Note. p values for test of null hypothesis that trend for each group is flat: English only, p = .886; bilingual, primarily English, p = .109; bilingual, primarily Spanish, p = .009; Spanish only, p = .026; p value for test of null hypothesis that gap trends are equal for all groups: p < .001.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

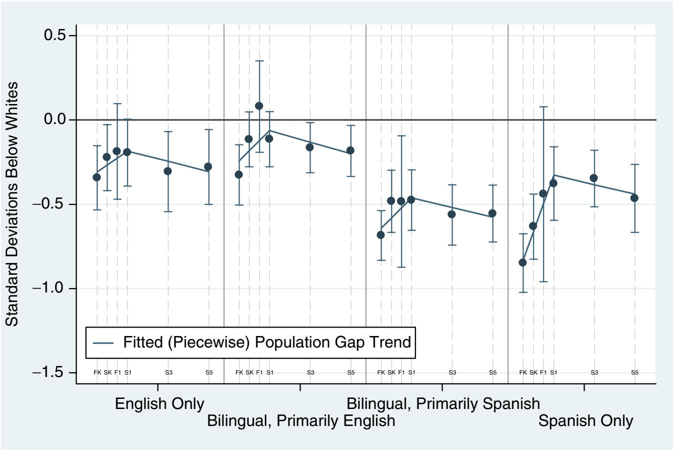

Figure 8. **Trends in estimated Hispanic-White reading gaps, by language used in home.**
Note. p values for test of null hypothesis that trend for each group is flat: English only, $p = .241$; bilingual, primarily English, $p = .516$; bilingual, primarily Spanish, $p = .202$; Spanish only, $p = .197$; p value for test of null hypothesis that gap trends are equal for all groups: $p < .001$.

A pattern similar to that shown in Figure 7 is evident in Table 3 above: Although non-English-proficient Hispanic students initially score considerably worse than English-proficient students at each wave, the non-English-proficient students gain more math skills from kindergarten to fifth grade relative to White students (half a standard deviation) than do the initially English-proficient Hispanic students (who gain only a sixth of a standard deviation). These gains, however, are not sufficient to make up the substantial initial gaps between the proficient and nonproficient students. In reading, initially nonproficient students in third and fifth grade lag by more than a standard deviation behind White students and by three quarters of a standard deviation behind Hispanic students who were proficient in English in kindergarten. Recall, however, that the large gaps between White students and Hispanic students who are not proficient in oral English at the start of kindergarten are likely not entirely due to differences in English proficiency at the start of schooling. As noted earlier, non-English-proficient students are disproportionately from low socioeconomic status homes, which likely contributes to their lower math scores. In analyses not shown here, we find that roughly a third to a half of the difference in fifth-grade reading scores between proficient and non-English-proficient Hispanic students can be accounted for by the lower socioeconomic status of the non-English-proficient students.

Achievement gaps by socioeconomic status quintile. Table 5 shows the distributions of socioeconomic status, by race/ethnicity and Hispanic

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Table 5

Proportion in Each Socioeconomic Quintile, by Race, Hispanic National Origin, and Mexican Immigrant Generational Status

	Socioeconomic Quintile				
	1 (*Low*)	2	3	4	5 (*High*)
White, not Hispanic, third+ generation	8.2	16.5	22.4	22.9	30.0
Hispanic, any race	35.8	24.5	17.7	13.4	8.6
Mexican origin	42.6	26.2	14.6	10.3	6.2
Mexican, first generation	66.9	26.3	4.9	0.0	1.9
Mexican, second generation	53.0	28.4	9.4	6.6	2.6
Mexican, third generation	17.6	22.5	26.5	19.4	14.0
Cuban origin	14.1	4.8	36.6	17.6	27.0
Puerto Rican origin	17.5	24.4	27.5	21.7	8.8
Central American origin	45.8	23.4	12.7	7.7	10.4
South American origin	9.9	41.1	14.2	15.3	19.5
Other Hispanic origin	35.7	15.2	25.0	13.4	10.8
National origin unknown	17.0	19.3	25.7	27.1	10.9
Black, not Hispanic, third+ generation	33.1	23.3	24.4	14.8	4.4
Asian, any generation	18.5	12.9	16.5	17.7	34.4
Other race, any generation	18.0	21.6	29.3	14.1	17.0
Total	18.3	19.3	21.9	19.2	21.4

Note. Percentages are weighted by the ECLS-K longitudinal weight C1_6FC0 and include only students in the longitudinal sample who have valid math scores in Wave 1 (i.e., those in the math column of Table 1). Percentages are not reported for Mexican–unknown generation group (owing to small sample size).

students' country of origin and Mexicans' immigrant generational status. Overall, Hispanic students come from families with much lower socioeconomic status than that of White students and similar to that of Black students. Among Hispanics, students of Central American and Mexican origin (particularly, first- and second-generation Mexican students) have the lowest average socioeconomic status. Roughly half these students (and more than three quarters of first-generation Mexican-origin students) come from families in the lowest socioeconomic status quintile. In contrast, Hispanic students of Cuban and South American origin have the highest socioeconomic status among Hispanic subgroups, showing a socioeconomic status distribution somewhat similar to that of Whites.

Figures 9 and 10 describe the Hispanic-White achievement gaps within each socioeconomic status quintile. Specifically, these figures show the average difference in math and reading scores between Hispanic and White students who are in the same socioeconomic quintile. Notably, at the start of kindergarten, Hispanic students score roughly a quarter to a third of a standard deviation lower in math and reading than do White students of the same socioeconomic status quintile.[20] By the spring of fifth grade, however, the gaps are smaller. In some cases—particularly among students in the lowest socioeconomic status quintile—Hispanic and White students' average

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

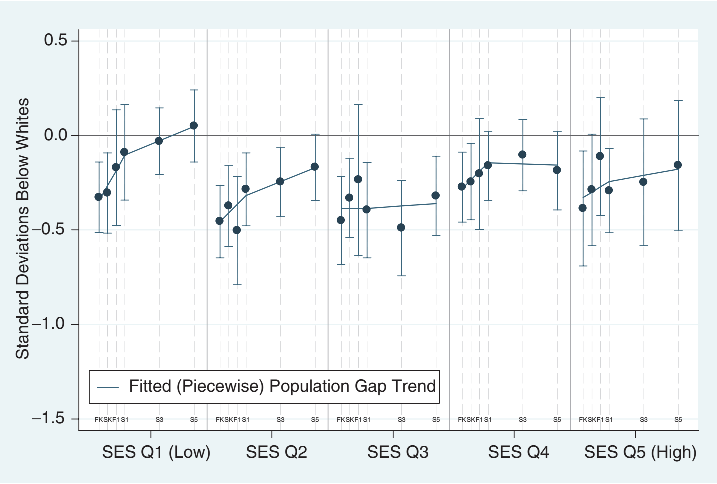

Figure 9. **Within-quintile trends in estimated Hispanic-White math gaps, by socioeconomic status.**
Note. p values for test of null hypothesis that trend for each group is flat: Q1, $p = .003$; Q2, $p = .049$; Q3, $p = .964$; Q4, $p = .093$; Q5, $p = .578$; p value for test of null hypothesis that gap trends are equal for all groups: $p = .060$.

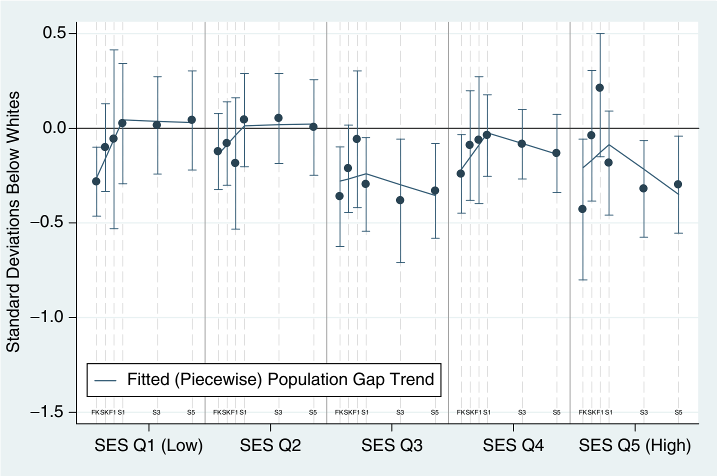

Figure 10. **Within-quintile trends in estimated Hispanic-White reading gaps, by socioeconomic status.**
Note. p values for test of null hypothesis that trend for each group is flat: Q1, $p = .239$; Q2, $p = .872$; Q3, $p = .590$; Q4, $p = .072$; Q5, $p = .660$; p value for test of null hypothesis that gap trends are equal for all groups: $p = .762$.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

scores are no different from one another. At higher socioeconomic status levels, Hispanic students still score slightly below White students, although the gaps are typically 0.10 to 0.20 standard deviations smaller than at the start of kindergarten.

Discussion

In this article, we provide a detailed descriptive analysis of the development of Hispanic-White achievement gaps in the elementary grades. We further previous research on Hispanic students' education by taking into account variations in these patterns among Hispanic subgroups and by focusing on early ages. Whereas older Hispanics' educational disadvantages have been relatively well documented, we know little about younger Hispanic children's schooling.

Five important findings emerge from this study. First, Hispanic students enter kindergarten with math and reading skills significantly lower than those of White students. In the fall of their kindergarten year, Hispanic students' math scores are three quarters of a standard deviation below those of White students. Among the 71% of Hispanic students who are proficient in oral English in the fall of kindergarten, their reading scores are a half standard deviation below those of White students. Second, unlike the Black-White test score gaps measured in the ECLS-K, which widen steadily from kindergarten through fifth grade, the Hispanic-White achievement gaps narrow during kindergarten and first grade (from 0.77 standard deviations to 0.56 in math and from 0.52 to 0.29 in reading). The Hispanic-White gaps change little in the years following first grade, however. By fifth grade, the math gap is still a half standard deviation, and the reading gap widens slightly, to three eighths of standard deviation.

Third, there is considerable variation in math and reading achievement patterns among Hispanic subgroups. Students of Mexican and Central American origins—particularly, students whose parents are immigrants—enter school with lower math and reading scores than do children of Cuban, South American, and other national origins and children of Hispanic parents born in the United States. Because students of Mexican and Central American origins are, on average, more socioeconomically disadvantaged than other Hispanic students and less likely to come from homes where English is spoken, it is possible that these patterns can be largely explained by socioeconomic and language differences among Hispanic subgroups. The lower socioeconomic status of Mexican and Central American students and first- and second-generation Mexican students may account for the lower math and reading skills of these students when they enter kindergarten.

Fourth, differences among Hispanic subgroups persist through fifth grade, although they become somewhat less pronounced. In fifth grade, Mexican-origin students score one quarter to one half a standard deviation lower than Cuban and Puerto Rican students on the ECLS-K tests of math and reading.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Fifth, those Hispanic subgroups with lower levels of math and reading skills at kindergarten entry show the most substantial narrowing of achievement gaps over time. By the end of fifth grade, math achievement gaps of Mexican students with foreign-born parents, Hispanic students from Spanish-speaking homes, and Hispanic students from the lowest quintile of socioeconomic status are a half standard deviation smaller than the corresponding gaps observed at kindergarten entry. In contrast, native-born Mexican students, Hispanic students living in English-speaking homes, and Hispanic students from the highest socioeconomic status quintiles show smaller gap reductions over time.

Although the ECLS-K data are not particularly well suited to investigate the causes of these patterns, it is worth considering the extent to which the patterns that we observe are consistent with the three major explanations for the gaps discussed earlier. First, some evidence is consistent with the argument that Hispanic-White differences in socioeconomic status are responsible for the gaps; the evidence of large gaps at kindergarten entry suggests that out-of-school factors are partly responsible for the gaps; and the gaps are largest, on average, for the most socioeconomically disadvantaged groups. Nonetheless, it is worth noting that first- and second-generation Mexican students are the most socioeconomically disadvantaged of the Hispanic subgroups that we examined, with average socioeconomic levels far below those of native-born Black students—for instance, 93% of first-generation Mexican students and 81% of second-generation Mexican students are in the bottom two socioeconomic status quintiles, compared to 53% of Black students (see Table 5). However, despite starting kindergarten with math and reading scores far below those of Black students and even further below those of White students, first- and second-generation Mexican students make substantial test score gains relative to Whites and Blacks during elementary school. By fifth grade, first- and second-generation Mexican students have average scores considerably higher than those of Black students and much closer to those of White students than they do in kindergarten (compare Figures 1 and 2 with Figures 5 and 6). These patterns cannot be accounted for by socioeconomic status differences among Mexican, Black, and White students.

Second, the evidence here is consistent with the hypothesis that at least part of the gains made by Hispanic students in the early grades are due to students' increased English acquisition (both oral and written)—which likely improves test performance on tests given in English (such as the reading test) and increases the opportunity for students to learn in schools where at least some, if not all, of the instruction is in English (Saunders & O'Brian, 2007). This evidence comes from several factors. The greatest relative gains in achievement are observed in groups with the lowest levels of home English use and oral English proficiency—that is, students of Mexican and Central American origin (particularly, those whose parents were born outside the United States and those who are from homes where English is not the predominant language). In addition, the narrowing of the

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

math and reading gaps primarily occurs in the first 2 years of schooling, when English acquisition is most rapid for these students. Moreover, Hispanic students not proficient in English at the start of kindergarten make much more rapid gains in math scores than do students proficient in English (see Table 3), suggesting that English acquisition may be a factor in their learning process (Saunders & O'Brian, 2007).

Finally, some of the rapid progress of Hispanic students in the first 2 years of schooling may be due to the use of instructional practices that are effective with English-language learners in the first years of schooling. The patterns of rapid gains in learning among groups with the lowest levels of initial English proficiency could result from the targeting of effective instructional practices on these students in the early elementary grades. Language support programs, for example, are more common in schools with large proportions of recent immigrants and students with low levels of English proficiency. To the extent that such practices are effective in narrowing the achievement gap for English learners, they may account for some of the rapid narrowing of the gaps in the early grades, when English proficiency is lowest for the students (in the ECLS-K sample). After first grade, instructional practices targeted to English learners may be less common and/or less effective in narrowing the gaps because, by this time, most students are reasonably proficient in oral English (in the ECLS-K sample, more than 80% of those not proficient in English in kindergarten are proficient by the end of first grade).

A final puzzle remains in the patterns reported here. Black and Hispanic students enter kindergarten with achievement levels equally low relative to White students. Moreover, the socioeconomic background of Black and Hispanic students (at least as measured by parental education, occupation, and income) are roughly equal at the start of kindergarten. Yet the achievement trajectories of Black and Hispanic students are strikingly different in the next 6 years: Black-White achievement gaps grow whereas Hispanic-White gaps narrow. Moreover, socioeconomic status explains virtually all of the Black-White gap at kindergarten entry (Fryer & Levitt, 2004) but much less of the Hispanic-White gap at the same age. However, by fifth grade, socioeconomic status no longer accounts for the entire Black-White gap (Fryer & Levitt, 2006), but it does for the majority of the Hispanic-White gap (see Figures 9 and 10). The reasons for this divergence between Black and Hispanic trajectories are beyond the scope of this article, although several explanations should be explored in future work.

First, the source of the initial achievement gaps at the start of kindergarten may differ for Black and Hispanic students. If Hispanic-White gaps are partly due to low English proficiency (of children and their parents), thereby limiting children's access to cognitively stimulating materials in English, and if Black-White gaps are primarily due to low socioeconomic status and its attendant family stressors, this may explain why Hispanic children show strong gains in the early grades (as they learn English) while Black students fall further behind.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Second, that Black and Hispanic students have similar socioeconomic backgrounds and cognitive skills at the start of schooling but different developmental trajectories after the start of schooling suggests that socioeconomic status is not the only factor that shapes achievement gaps. Differences in the average quality of schools attended by Hispanic, Black, and White students may play an important role in the development of achievement disparities (for Blacks) or their amelioration (for Hispanics). That Hispanic-White gaps narrow following the start of formal schooling suggests that schooling has the potential to remedy at least part of initial achievement gaps. Future research might identify the features of schooling most strongly associated with the reduction of Hispanic-White achievement gaps in the early grades and so test ways of improving schools to further narrow the gaps.

Appendix
Estimation of Achievement Gaps

Estimating Wave-Specific Gap Measures

Let θ_{ist} denote the observed test score in subject s at wave t for student i. Let G be a vector of dummy variables indicating group or subgroup (with Whites being the omitted group). Let T_{it} indicate the assessment date for student i at wave t. We estimate three types of gaps using the θ scores.

Mean θ score differences. For each subject s and wave t, we estimate the vector of gaps Δ_{st} between each Hispanic subgroup and Whites by fitting the following model (via weighted least squares, using the ECLS-K child longitudinal weight C1_6FC0):

$$\theta_{ist} = \beta_0 + \beta_1 T_{it} + \mathrm{G}_{ist}\Delta_{st} + \varepsilon_{ist}. \qquad \text{A1}$$

The vector $\widehat{\Delta}_{st}$ contains the estimated gaps in subject s at wave t between each group (or subgroup) and Whites. We compute standard errors adjusted for the complex sampling design of the ECLS-K, using the *svy* commands in Stata 9.1 (Stata Corporation, College Station, Texas).

Standardized mean differences. We first estimate each student's true θ score in each subject s and wave t, given an assumed reliability r (assumed reliabilities ranging from 0.7 to 1.0). Assuming classical measurement error, we estimate the true value of θ as

$$\theta^*_{ist} = (1-\sqrt{r})\bar{\theta}^i_{st} + (\sqrt{r})\theta_{ist}, \qquad \text{A2}$$

(continued)

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Appendix (continued)

where $\bar{\theta}^{i}_{st}$ is the mean θ for student *i*'s race/ethnic group (White, Black, Hispanic, Asian, other) in subject *s* at wave *t*. The resulting θ^{*}_{ist} distribution will have smaller within-group variance than that of θ_{ist} but the same group means (essentially, we are subtracting the proportion of the variance assumed to be due to measurement error).

We next compute the assessment-date adjusted pooled within-race/ethnic group standard deviation of test scores at each wave and test subject. For both Hispanic and White students separately, we fit (via weighted least squares) a set of models (for both math and reading and at each of the six waves) of the form

$$\theta^{*}_{ist} = \beta_0 + \beta_1 T_{it} + \varepsilon_{ist}. \qquad \text{A3}$$

From each model, we obtain, for Hispanics and for Whites, an estimate of the standard deviation of test score *s* at wave *t* (adjusted for measurement error and assessment date). We compute the estimated pooled standard deviation for each test subject at each wave, $\hat{\sigma}^{*}_{st}$, as the square root of the average of the squares of the subject- and wave-specific standard deviations (Hispanic and White). We next standardize the test scores by de-meaning them and dividing by the estimated pooled standard deviation:

$$\theta'^{*}_{ist} = \frac{\theta^{*}_{ist} - \bar{\theta}^{*}_{st}}{\hat{\sigma}^{*}_{st}} \qquad \text{A4}$$

To estimate the standardized gaps, we refit Model A1 using the reliability-adjusted standardized test score:

$$\theta'^{*}_{ist} = \beta'^{*}_{0} + \beta'^{*}_{1} T_{it} + \mathrm{G}_{ist}\Delta'^{*}_{st} + \varepsilon'^{*}_{ist} \qquad \text{A5}$$

As above, the vector $\widehat{\Delta}'^{*}_{st}$ contains the estimated gaps in subject *s* at wave *t* between each group (or subgroup) and Whites, adjusted for measurement error, although here all between-group differences in test scores are expressed in terms of measurement-error-adjusted pooled standard deviations.

Metric-free gap measures. Let $F_{hst}(\theta)$ and $F_{wst}(\theta)$ indicate the cumulative density functions of the Hispanic and White test score distributions (corrected for measurement error) in subject *s* at wave *t*, respectively. Define $H_{st}[F_{hst}(\theta)] = F_{wst}(\theta)$; that is, let H_{st} be the function that maps percentiles of the Hispanic distribution of scores in subject *s* at wave *t* onto percentiles of the corresponding White distribution. Then the probability that a randomly chosen Hispanic student has a score in subject *s* at wave *t* that is higher than that of a randomly chosen student is given by

(continued)

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Appendix (continued)

$$P_{h>w(st)} = \int_0^1 H_{st}(x)dx. \quad \text{A6}$$

We plot H_{st} and compute the integral in Equation A6 numerically via interpolation. $P_{g>w(st)}$ is similarly computed for any Hispanic subgroup *g*. Finally, Ho and Haertel's pseudo-effect size (2006) is computed as

$$\delta_{st} = \sqrt{2}\Phi^{-1}\left(P_{h>w(st)}\right), \quad \text{A7}$$

where Φ^{-1} is the inverse cumulative normal density function.

Fitting Achievement Trend Lines

Model A5 yields estimated standardized gaps and their standard errors for group *g* in subject *s* at wave *t*. For each group and subject, we have six such estimates, one at each wave. We fit piecewise trend lines through each set of estimates (see Figures 1–10), with a kink at the spring of first grade, weighting each estimate by the inverse of its sampling variance $\left(\left[se\left(\hat{\Delta}'^{*}_{gst}\right)\right]^{-2}\right)$ in the regression. Experimentation with alternate specifications, including linear and quadratic trend lines, as well as piecewise trend lines with a kink at another point, did not provide as good a fit, on average, as the piecewise specification that we chose.

Notes

The research reported here was funded by the Carnegie Scholars Program of the Carnegie Corporation of New York. Additional research support was provided by the National Task Force on Early Childhood Education for Hispanics and the AERA Research Grants Program. We thank Eugene García, L. Scott Miller, and three anonymous reviewers for helpful feedback and suggestions and Joe Robinson and Tara Beteille for exceedingly good research assistance. All errors remain our own.

[1]Throughout the article, we use the terms *White*, *Black*, and *Asian* to denote non-Hispanic members of each group.

[2]Recent nationally representative studies with sizable Hispanic student samples include the National Assessment of Educational Progress (NAEP; 4th, 8th, and 12th grade), the National Education Longitudinal Study (8th–12th grade), the National Longitudinal Study of Adolescent Health (7th–12th grade), and the Education Longitudinal Study (10th–12th grade). The Early Childhood Longitudinal Study–Kindergarten Class of 1998–99 (ECLS-K), which we use in this article, is the only recent nationally representative study with a large Hispanic sample of elementary school students (kindergarten–5th grade).

[3]Here we adopt the common definitions used in describing immigrant generational status: First-generation immigrants are those born outside the United States to nonnative parents; second-generation immigrants are those born in the United States to nonnative parents; and third-generation immigrants are those born to parents who were born in the United States.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

[4]The estimates are based on our calculations from NAEP data (obtained at http://nces.ed.gov/nationsreportcard/nde/). Detailed tabulations are available upon request. Some caution is warranted when comparing Hispanic-White achievement gaps across cohorts or grades. Given the dramatic increase of the Hispanic student population since the early 1990s, the composition of Hispanic students tested in different grades and different years may vary somewhat, and so any comparison across cohorts or grades may be complicated by the changing composition of the Hispanic student population. In addition, changes in the content of the NAEP assessments may confound estimates of changes in the size of achievement gaps, although gap estimates are arguably less sensitive to changes in the content of the test than are estimates of mean score differences (discussion of measurement issues is discussed later in the article). Finally, NAEP exclusion rates (i.e., the rates at which students identified as English learners were excluded from the sample) vary somewhat across years, grades, and test subjects, further complicating comparison of the trends among states. Nonetheless, because NAEP exclusion rates were generally higher in the early years of the assessment than in recent years, the apparent narrowing trend in its gaps likely underestimates the true narrowing trend.

[5]Although other longitudinal studies report Hispanic-White achievement gap estimates, we do not discuss these for a variety of reasons: Some are based on small local samples (e.g., Bali & Alvarez, 2003) and others on nonrepresentative national samples; for example, Phillips (2000) estimates Mexican-White gaps using data from the National Longitudinal Survey of Youth, which includes a sample of children born to women who were 14–21 years old and living in the United States in 1979—a population that excludes first-generation children and children of parents who immigrated since 1979 and that is overrepresentative of children born to young mothers. Furthermore, some report gaps in ways not comparable to other studies (e.g., Murnane, Willett, Bub, & McCartney, 2006).

[6]Data patterns among the first-grade cohort in the Prospects study are somewhat perplexing: The math gaps widen sharply in the first and second grade, then narrow sharply in the third; the reading gaps, however, narrow in the first grade, widen in second, and then narrow again in third. Given the general consistency among the other studies regarding a gradual narrowing, the patterns in the Prospects first-grade cohort data are likely due to sampling variation, or they are an artifact of the tests used. Borman, Stringfield, and Rachuba (2000) do not report standard errors for their estimates; so, we cannot assess the extent to which sampling variation may account for the inconsistency.

[7]About a third of the Hispanic students in the ECLS-K study are missing information on national/regional origin. Most of these cases occur because the student left the study before first grade, when the parental birthplace questions were included in the survey. However, we focus our analysis on the subsample of Hispanic students who were assessed through fifth grade, almost all of whom have data on immigrant generation and national/regional origin.

[8]Roughly 23% of the total ECLS-K sample (roughly 26% of Hispanic students) cannot be unambiguously categorized as first-, second-, or third-generation. Most of this missing data (17% of the total sample) is a result of the fact that the kindergarten parent survey did not ask about the parent's country of birth (it was asked in the first- and third-grade surveys); so, we cannot distinguish second- from third-generation students among those who left the sample before the spring of first grade (about a fifth of the total sample). However, because we included in our analyses only students present in the study through fifth grade, few members of our analytic sample are missing generational status data.

[9]Students took the math assessment in English if either the school records or the students' teachers reported that they were from homes where English was the primary language or if they passed the English Oral Language Development Scale assessment given by ECLS-K assessors. Hispanic students otherwise took the test in Spanish. Some evidence suggests that the language screening may not have identified all non-English-proficient Hispanic students: Rumberger and Tran (2006) report that 16% of ECLS-K kindergarten students from homes where parents reported Spanish as the dominant home language were not identified as being from Spanish-speaking homes by their

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

schools or their teachers (and so were not given the English Oral Language Development Scale screening test). However, only half of Hispanic students were from homes where parents reported Spanish as the primary language, thus implying that screening failures may have misclassified 8% of Hispanic students, at most (and probably fewer). The math test did not require students to read or write in either Spanish or English; the assessor read the question and all responses aloud to the child. The reading tests did, of course, require students to read or recognize letters and words in English, but they did not require the students to write in English until the third- and fifth-grade assessments; in the kindergarten and first-grade assessments, students responded orally or by pointing (Tourangeau, Nord, Lê, Pollack, & Atkins-Burnett, 2006). There were no other special accommodations provided for non-English-proficient students.

[10]It is important to consider the extent to which estimates of the gap may be influenced by what is commonly known as *test bias*—that is, bias that might result if a test does not measure the desired dimension of cognitive skill equally well for Hispanic and non-Hispanic White students, unconfounded by language or cultural knowledge. In the ECLS-K data, several facts lead us to believe that the tests do not suffer from such biases. First, the math tests were administered orally, in Spanish, to Hispanic students not proficient in oral English, thereby ensuring that math test performance was not dependent on students' oral English proficiency nor on literacy skills in English or Spanish—although Rumberger and Tran (2006) found that 16% of students from homes where Spanish was the primary language were not given the oral English proficiency assessment in kindergarten, meaning that it is possible that some Hispanic students who were not proficient in English were mistakenly assessed in English rather than in Spanish. Second, the reading tests were not administered to students not proficient in oral English. This limits the sample and generalizability of the reading estimates, but it ensures that the estimated reading gaps are unbiased with regard to the population of English-proficient students. Third, the developers of the ECLS-K assessments performed a range of differential item functioning tests on the assessment items to ensure that there was no item-level bias in the cognitive assessments (for detail, see Pollack, Narajian, Rock, Atkins-Burnett, & Hausken, 2005).

[11]At Wave 3 of the ECLS-K study (fall of first grade), only a 30% random subsample of the full ECLS-K sample was assessed in math and reading. Because missing test score data at this wave are, by design, completely missing at random, we include students who were not assessed in Wave 3 in our analytic sample as long as they were present at all other waves.

[12]Although we use the longitudinal sample in the analyses reported here, our results do not substantially differ from results based on all available students at each wave. In other words, sample attrition does not appear to be systematically related to achievement once the child longitudinal weights are applied. However, for some of the smallest subgroups, attrition of a small number of students does appear to affect our gap estimates. We report the gap trends for a constant sample to avoid confounding the gap trends with attrition noise.

[13]When weighted by the ECLS-K weight C1_6FC0, Hispanics make up 20% of our sample, whereas estimates based on the 1999 Current Population Survey suggest that 17%-18% of U.S. kindergartners were Hispanic (see Table 2 at http://www.census.gov/population/www/socdemo/school/p20-533.html). Among Hispanics, 57% of our sample is Mexican (compared to 59% nationwide); 6%, Puerto Rican (10%); 8%, Central American or Dominican (7%); 2%, Cuban (4%); and 5%, South American (4% nationwide). National figures are from Ramirez (2004) and are based on the full Hispanic population in 2000, whereas our sample represents the fall 1998 kindergarten population.

[14]More important, we do not rely on the published ECLS-K "scale scores," which are estimates of the number of questions that a student would have gotten right had she or he been administered all items on the full K–5 test (no student was administered every item at any wave; the adaptive nature of the test meant that students were asked only the items that were appropriate to their skill levels). The scale score is an arbitrary and highly nonlinear monotonic transformation of the θ score, meaning that conclusions about the

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

patterns of Hispanic-White gaps substantially depend on the transformation used (which is a function of the distribution of item difficulties on the test), a point that Reardon (2008) empirically demonstrates with regard to Black-White gaps. To see this, consider two individuals who start kindergarten with different θ scores and who experience the same increase in θ over time. Because the scale score metric is a nonlinear transformation of the θ metric, the two students will not, in general, experience the same increase in their scale scores. Thus, inferences about changes in the size of the achievement gap between the two students will differ, depending on which metric is used. Although the θ scores have not been included in the data files provided by NCES, files containing the θ scores are available from the NCES upon request (or can be constructed from the published data using the approach shown in Reardon, 2008).

[15]We use the term *interval-scaled* to mean that observed test scores are a linear function of true (unobservable) cognitive skill (plus random mean-zero measurement error), as opposed to the weaker condition that test scores are a monotonically increasing function of cognitive skill (plus error).

[16]The standardized gap measures are computed by dividing the mean score difference by the pooled standard deviation of test scores. If test scores contain measurement error, the estimated pooled standard deviation will typically overestimate the within-group standard deviation in true scores, thereby leading to a bias toward zero in the estimated standardized gaps.

[17]Estimated internal item-level reliabilities of the ECLS-K scores range from .89 to .96 (Pollack, Narajian, et al., 2005), and the test-retest reliability of the ECLS-K tests is likely somewhere between .75 and .95 (based on estimates of test-retest reliability of IQ and school readiness tests for children aged 5–10; see Rock & Stenner, 2005). Thus, the overall reliability of the test scores is likely between .70 and .90, which means that the estimated standard deviation of true scores is likely 5% to 15% smaller than the estimated standard deviation of the observed scores. As a result, estimates of the between-group gaps that assume no measurement error will underestimate the magnitude of the true gaps by 5% to 15%. To adjust the gaps for measurement error, we assume classical measurement error in the test. We use an assumed test reliability to shrink each student's observed test score toward his or her group (Hispanic or White) mean score (see appendix). We then compute the Hispanic-White gaps using these shrunken test scores. For discussion of the meaning of the different types of reliability and its attenuation effects, see Henson (2001) and Nunnally and Bernstein (1994).

[18]Standard errors for the metric-free measures cannot be computed analytically but could be computed by bootstrapping. Bootstrapped standard errors for the pseudo-effect sizes would be similar to those for the standardized difference measures, given the approximate normality of the θ distributions. We have not gone to the trouble of computing bootstrapped standard errors, however, because we are not interested in making statistical comparisons here.

[19]Our Hispanic-White standardized gap estimates differ slightly from those reported in Fryer and Levitt (2006) because we use different sample selection criteria, adjust our estimates for the date of assessment, and use the pooled standard deviation (adjusted by assessment date) rather than the sample standard deviation at each wave.

[20]In analyses not shown here, we observe a clear socioeconomic gradient in achievement patterns of Hispanic students. Hispanic students in the lowest socioeconomic status quintiles start kindergarten with math and reading scores significantly below those of the average White student, although low socioeconomic status Hispanic students make the largest gains relative to White students in kindergarten and first grade.

References

Bali, V. A., & Alvarez, R. M. (2003). Schools and educational outcomes: What causes the "race gap" in student test scores? *Social Science Quarterly, 84*(3), 485–507.

Ballou, D. (2008, April). *Test scaling and value-added measurement*. Paper presented at the National Conference on Value-Added Modeling, Madison, WI.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Borman, G. D., Stringfield, S., & Rachuba, L. (2000). *Advancing minority high achievement: National trends and promising programs and practices*. Baltimore: Johns Hopkins University, Center for Social Organization of Schools.

Cameron, S., & Heckman, J. (2001). The dynamics of educational attainment for Black, Hispanic, and White males. *Journal of Political Economy, 109*, 455–499.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). *The academic achievement gap in Grades 3 to 8* (Working Paper No. 12207). Cambridge, MA: National Bureau of Economic Research.

Crawford, J. (1997). *Best evidence: Research foundations of the Bilingual Education Act*. Washington, DC: National Clearinghouse for Bilingual Education.

DeNavas-Walt, C., Proctor, B. D., & Mills, R. J. (2004). *Income, poverty, and health insurance coverage in the United States: 2003* (Report No. P60–266). Washington, DC: U.S. Census Bureau.

Duncan, G. J., & Magnuson, K. A. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps? *The Future of Children, 15*(1), 35–54.

Fry, R. (2004). *Hispanic youth finishing college: The role of selective pathways*. Washington, DC: Pew Hispanic Center.

Fryer, R. G., & Levitt, S. D. (2004). Understanding the Black-White test score gap in the first two years of school. *Review of Economics and Statistics, 86*(2), 447–464.

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.

Galindo, C., & Reardon, S. F. (2006). *Latino students' educational experiences and opportunities during kindergarten*. Tempe, AZ: National Task Force on Early Childhood Education for Hispanics.

Gándara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English learners in California schools: Unequal resources, unequal outcomes. *Educational Policy Analysis Archives, 11*(36). Retrieved February 11, 2009, from http://epaa.asu.edu/epaa/v11n36/

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, DC: Brookings Institution Press.

Guzman, B. (2001). *The Hispanic population: Census 2000 brief* (Current Population Reports No. C2KBR/01–3). Washington, DC: U.S. Census Bureau. Retrieved February 10, 2009, from http://www.census.gov/prod/2001pubs/c2kbr01–3.pdf

Hakimzadeh, S., & Cohn, D. V. (2007). *English usage among Hispanics in the United States*. Washington, DC: Pew Hispanic Center.

Hanushek, E. A., & Rivkin, S. G. (2006). *School quality and the Black-White achievement gap*. Unpublished manuscript.

Hao, L., & Bonstead-Bruns, M. (1998). Parent-child differences in educational expectations and the academic achievement of immigrant and native students. *Sociology of Education, 71*, 175–198.

Hedges, L. V., & Nowell, A. (1999). Changes in the Black-White gap in achievement test scores. *Sociology of Education, 72*(2), 111–135.

Henderson, R., & Landesman, E. (1992). *Mathematics and middle school students of Mexican descent: The effect of thematically integrated instruction*. Santa Cruz: University of California, Center for Research on Education, Diversity, and Excellence.

Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counseling and Development, 34*, 177–189.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Hernandez, D. J. (2006). *Young Hispanic children in the U.S.: A demographic portrait based on Census 2000*. New York: National Task Force on Early Childhood Education for Hispanics.

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: Center for the Study of Evaluation.

Kao, G., & Thompson, J. (2003). Race and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology, 29*, 417–442.

Kaufman, P., Alt, M. N., & Chapman, C. D. (2001). *Dropout rates in the United States: 2000* (NCES Report No. 2002–114). Washington, DC: U.S. Department of Education.

Kohler, A., & Lazarín, M. (2007). *Hispanic education in the United States*. Washington, DC: National Council of La Raza.

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.

Lichter, D. T., Qian, Z., & Crowley, M. L. (2005). *Poverty and economic polarization among America's minority and immigrant children*. Ann Arbor, MI: National Poverty Center.

LoGerfo, L., Nichols, A., & Reardon, S. F. (2006). *Achievement gains in elementary and high school*. Washington, DC: Urban Institute.

Lord, F. M., & Novick, M. R. (1968). *Statistical theory of mental test scores*. Reading, MA: Addison-Wesley.

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, 2006,* 97–135.

National Center for Education Statistics. (2001). *ECLS-K, base year public-use data file and user's manual* (NCES Report No. 2001–029). Washington, DC: U.S. Department of Education.

National Center for Education Statistics. (2002a). *Common core of data (CCD) public elementary and secondary school universe data*. Retrieved February 10, 2009 from http://www.nces.ed.gov/ccd/pubschuniv.html

National Center for Education Statistics. (2002b). *User's manual for the ECLS-K first grade public-use data files and electronic codebook* (NCES Report No. 2002–135). Washington, DC: U.S. Department of Education.

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.

Neut, A. (2006). Hispanics in the US: Mobility within the US. *Monitor Hispano*. Retrieved September 25, 2008, from http://mediaserver.fxstreet.com/Reports/9c1b6d93–dd98–4fa6–b31a-4e881a171ed3/fab08094–62cb-4cb2–9fac-6028b70ac984.pdf

Nunnally, J. C., & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Orfield, G., & Yun, J. T. (1999). *Resegregation in American schools*. Retrieved from http://www.law.harvard.edu/groups/civilrights/publications/resegregation99.htm

Padilla, A., & Gonzalez, R. (2001). Academic performance of immigrant and U.S. born Mexican heritage students: Effects of schooling in Mexico and bilingual English language instruction. *American Educational Research Journal, 38*(3), 727–742.

Padron, Y. N., Waxman, H. C., & Rivera, H. H. (2002). *Educating Hispanic students: Obstacles and avenues to improve academic achievement* (Report

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

No. EPR08). Santa Cruz: University of California, Center for Research on Education, Diversity, and Excellence.

Perreira, K., Harris, K., & Lee, D. (2006). Making it in America: High school completion by immigrant and native youth. *Demography, 43*, 511–536.

Phillips, M. (2000). Understanding ethnic differences in academic achievement: Empirical lessons from national data. In D. W. Grissmer & J. M. Ross (Eds.), *Analytic issues in the assessment of student achievement* (pp. 103–132). Washington, DC: National Center for Education Statistics.

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, DC: Brookings Institution Press.

Planty, M., Hussar, W., Snyder, T., Provasnik, S., Kena, G., Dinkes, R., et al. (2008). *The condition of education, 2008* (NCES Report No. 2008–031). Washington, DC: U.S. Department of Education.

Pollack, J. M., Narajian, M., Rock, D. A., Atkins-Burnett, S., & Hausken, E. G. (2005). *Early Childhood Longitudinal Study–Kindergarten Class of 1998–99 (ECLS-K), psychometric report for the fifth grade* (NCES Report No. 2006–036). Washington, DC: U.S. Department of Education.

Pollack, J. M., Rock, D. A., Weiss, M. J., Atkins-Burnett, S., Tourangeau, K., West, J., et al. (2005). *Early Childhood Longitudinal Study–Kindergarten Class of 1998–99 (ECLS-K), psychometric report for the third grade* (NCES Report No. 2005–062). Washington, DC: U.S. Department of Education.

Portes, A., & Schauffler, R. (1996). Language and the second generation: Bilingualism yesterday and today. In A. Portes (Ed.), *The new second generation* (pp. 8–29). New York: Russell Sage Foundation.

Ramirez, R. R. (2004). *We the people: Hispanics in the United States* (Report No. CENSR-18). Washington, DC: U.S. Census Bureau.

Reardon, S. F. (2003). *Sources of educational inequality: The growth of racial/ethnic and socioeconomic test score gaps in kindergarten and first grade* (Working Paper No. 03–05R). University Park: Pennsylvania State University, Population Research Institute.

Reardon, S. F. (2008). *Thirteen ways of looking at the Black-White test score gap* (Working Paper No. 2008-08). Stanford, CA: Stanford University, Institute for Research on Educational Policy and Practice.

Reardon, S. F., & Galindo, C. (2006). *Patterns of Hispanic students' math and English literacy test scores in the early elementary grades.* New York: National Task Force on Early Childhood Education for Hispanics.

Reardon, S. F., & Robinson, J. (2007). Patterns and trends in racial/ethnic and socioeconomic academic achievement gaps. In H. Ladd & E. Fiske (Eds.), *Handbook of research in education finance and policy* (pp. 497–516). New York: Routledge.

Rock, D. A., & Pollack, J. M. (2002). *Early Childhood Longitudinal Study–Kindergarten Class of 1998–99 (ECLS-K), psychometric report for kindergarten through first grade* (NCES Report No. 2002–05). Washington, DC: U.S. Department of Education.

Rock, D. A., & Stenner, A. J. (2005). Assessment issues in the testing of children at school entry. *The Future of Children, 15*(1), 15–34.

Rosenthal, A., Baker, K., & Ginsburg, A. (1983). The effects of English background on achievement levels and learning outcomes among elementary school students. *Sociology of Education, 56*(4), 157–169.

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011

Rothstein, R. (2004). *Class and schools. Using social, economic, and educational reform to close the Black-White achievement gap*. Washington, DC: Economic Policy Institute.

Rumberger, R. W., & Arellano, B. (2004). *Understanding and addressing the California Latino achievement gap in early elementary school* (Report No. 2004–01). Santa Barbara, CA: University of California Linguistic Minority Research Institute.

Rumberger, R. W., & Tran, L. (2006). *Preschool participation and the cognitive and social development of language-minority students* (CSE Technical Report No. 674). Los Angeles: University of California, Center for the Study of Evaluation.

Saunders, W. M., & O'Brian, G. (2007). Oral language. In F. Genesee, K. Lindholm-Leary, W. M. Saunders, & D. Christian (Eds.), *Educating English language learners: A synthesis of research evidence* (2nd ed., pp. 14–63). Cambridge, UK: Cambridge University Press.

Schmid, C. (2001). Educational achievement, language minority and the new second generation. *Sociology of Education, 73*(Extra issue), 71–87.

Schneider, B., Martinez, S., & Owens, A. (2006). Barriers to educational opportunities for Hispanics in the United States. In M. Tienda & F. Mitchell (Eds.), *Hispanics and the future of America* (pp. 179–227). Washington, DC: National Academies Press.

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.

Smith, J. P. (2006). Immigrants and their schooling. In E. A. Hanushek & F. Welch (Eds.), *Handbook of economics of education* (Vol. 1, pp. 155–187). New York: Elsevier.

Stoops, N. (2004). *Educational attainment in the United States: 2003*. Washington, DC: U.S. Census Bureau.

Thomas, W. P., & Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement*. Santa Cruz: University of California, Center for Research on Education, Diversity, and Excellence.

Tourangeau, K., Nord, C., Lê, T., Pollack, J. M., & Atkins-Burnett, S. (2006). *Early Childhood Longitudinal Study–Kindergarten Class of 1998–99 (ECLS-K), combined user's manual for the ECLS-K fifth-grade data files and electronic codebooks* (NCES Report No. 2006–032). Washington, DC: U.S. Department of Education.

Van Hook, J., & Balistreri, K. S. (2002). Diversity and change in the institutional context of immigrant adaptation: California schools 1985–2000. *Demography, 39*(4), 639–654.

Winsler, A., Diaz, R., Espinosa, L., & Rodriguez, J. (1999). When learning a second language does not mean losing the first: Bilingual language development in low-income, Spanish-speaking children attending bilingual preschool. *Child Development, 70*(2), 349–362.

Zill, N., Collins, M., West, J., & Hausken, E. G. (1995). *Approaching kindergarten: A look at preschoolers in the United States* (NCES Report No. 95–280). Washington, DC: U.S. Department of Education.

Manuscript received February 11, 2008
Final revision received January 8, 2009
Accepted January 13, 2009

Downloaded from http://aerj.aera.net at Stanford University Libraries on February 15, 2011