FEATURE ARTICLES

Science Achievement Gaps Begin Very Early, Persist, and Are Largely Explained by Modifiable Factors

Paul L. Morgan[1], George Farkas[2], Marianne M. Hillemeier[1], and Steve Maczuga[1]

We examined the age of onset, over-time dynamics, and mechanisms underlying science achievement gaps in U.S. elementary and middle schools. To do so, we estimated multilevel growth models that included as predictors children's own general knowledge, reading and mathematics achievement, behavioral self-regulation, sociodemographics, other child- and family-level characteristics (e.g., parenting quality), and school-level characteristics (e.g., racial, ethnic, and economic composition; school academic climate). Analyses of a longitudinal sample of 7,757 children indicated large gaps in general knowledge already evident at kindergarten entry. Kindergarten general knowledge was the strongest predictor of first-grade general knowledge, which in turn was the strongest predictor of children's science achievement from third to eighth grade. Large science achievement gaps were evident when science achievement measures first became available in third grade. These gaps persisted until at least the end of eighth grade. Most or all of the observed science achievement gaps were explained by the study's many predictors. Efforts to address science achievement gaps in the United States likely require intensified early intervention efforts, particularly those delivered before the primary grades. If unaddressed, science achievement gaps emerge by kindergarten and continue until at least the end of eighth grade.

Keywords: achievement gaps; at-risk students; early childhood; growth trajectories; hierarchical linear modeling; longitudinal; minorities; poverty; racial/ethnic minorities; regression analyses; science achievement; science education; secondary data analysis; socioeconomic status; survey research

Very large science achievement gaps are present in U.S. elementary and middle schools (e.g., U.S. Department of Education, 2000). For example, national estimates indicate that the 50th-percentile science scores of eighth graders who are Black approximate the 10th-percentile scores of eighth graders who are White (U.S. Department of Education, 2015). The 50th-percentile score of those who are English language learners (ELL) are lower than the 10th-percentile score of those who are non-ELL. The 50th-percentile science score of eighth-grade children eligible for free lunch are below the 25th-percentile score for those not eligible. These achievement gaps themselves likely result from opportunity gaps. For example, and due to historical as well as ongoing socioeconomic inequities (Ladson-Billings, 2006), children who are from traditionally marginalized groups attending U.S. schools are more likely to be provided with lower-quality educational experiences, including attending underfunded and lower-resourced schools where they are often taught by inexperienced teachers and/or those with lower expectations for the children's achievement (Flores, 2007). Children

from these groups also experience many opportunity gaps prior to school entry (Morsy & Rotherstein, 2015), including lower access to high-quality child care and preschools and the resulting informal learning opportunities and achievement gains that these settings can provide (Cascio & Schanzenbach, 2013; Hillemeier, Morgan, Farkas, & Maczuga, 2013; Ramani & Siegler, 2008).

Unfortunately, and as the United States experiences greater income inequality, science achievement gaps may be experienced by progressively larger percentages of U.S. schoolchildren (Gould, Mishel, & Shierholz, 2013; U.S. Census Bureau, 2012a, 2012b). Children with low levels of science achievement may be less able as adults to understand public policy issues necessitating ever-greater scientific literacy and reasoning (e.g., climate change, hydraulic fracturing, genetic engineering) as well as experience lower employment and prosperity (National Academy

[1]Pennsylvania State University, University Park, PA
[2]University of California, Irvine, CA

of Sciences, National Academy of Engineering, & Institute of Medicine [NASNAEIM], 2010, 2011).

Why these science achievement gaps occur is poorly understood. Few studies have investigated the age of onset, over-time stability, and mechanisms explaining these gaps (Bacharach, Baumeister, & Furr, 2003; Muller, Stage, & Kinzie, 2001; Von Secker, 2004). To date, the extant work has mostly analyzed science achievement in older (e.g., high school) student populations (e.g., Byrnes & Miller, 2007; Ma & Wilkins, 2002; Muller et al., 2001; Reynolds & Walberg, 1992; J. Wang & Staver, 2001; Young, Reynolds, & Walberg, 1996), at particular time points rather than longitudinally (Liu & Whitford, 2010; Maerten-Rivera, Myers, Lee, & Penfield, 2010; Von Secker, 2004), or using non-U.S. samples (Li, Onaga, Shen, & Chiou, 2009). Byrnes and Miller (2007) identified only 12 longitudinal studies of science achievement published since 1992. Of these 12 studies, the number of factors per study used as predictors of science achievement averaged only five, with a maximum of eight. Thus, rigorous multivariate analyses identifying risk factors for lower initial and over-time growth in science achievement have yet to be conducted. As stated by Byrnes and Miller (2007), "there is no way to tell the difference between important, authentic predictors and relatively minor or even spurious predictors" of U.S. schoolchildren's science achievement (p. 600). For example, none of the five studies reporting on factors associated specifically with racial and ethnic science achievement gaps used a longitudinal sample representative of the U.S. population of elementary and middle school students and incorporated extensive statistical control for confounding factors (i.e., Bacharach et al., 2003; Delen & Bulut, 2011; O. Lee, Deaktor, Hart, Cuevas, & Enders, 2005; Von Secker, 2004; Von Secker & Lissitz, 1999). The lack of empirical knowledge of the age of onset and over-time trajectories of science achievement gaps in the United States limits efforts by education researchers, policymakers, and practitioners to address these gaps, particularly as disproportionately experienced by racial, ethnic, and language minorities as well as low-income children.

Possible Dynamics of Science Achievement Gaps

Theoretically, the age of onset and over-time dynamics of science achievement gaps might follow one of three possible trajectories. The first possibility is that these gaps occur at early ages and are then largely stable (e.g., Reardon, 2011; Scarborough, 1998). Children's earlier levels of achievement are known to be highly predictive of later levels of achievement (Chatterji, 2006; Duncan et al., 2007; Morgan, Farkas, & Wu, 2009). The science achievement gaps might begin to occur as some groups of children tend to experience fewer informal opportunities to learn about science in their homes and preschools from birth to kindergarten entry (Sackes, Trundle, Bell, & O'Connell, 2011). These gaps might then remain stable as any subsequent achievement gains attributable to efforts by the children's teachers and schools remain relatively uniform across time and, for those who are at risk, are insufficient to close the initial gaps (McCoach, O'Connell, Reis, & Levitt, 2006).

A second possibility, often termed a cumulative advantage growth trajectory or Matthew effect, is that some children

initially experience low science achievement as well as increasing science achievement gaps (Morgan, Farkas, & Wu, 2011; Stanovich, 1986; Strand, 2014; Walberg & Tsai, 1983). Science achievement gaps might initially occur for children entering school with relatively less knowledge about the physical and natural sciences as a result of experiencing fewer informal learning opportunities prior to school entry. Over time, the children's knowledge about science might begin to lag further behind that of their peers. This might occur as a result of attending lower-quality and lower-resourced schools where the science (as well as mathematics and reading) instruction is of lower quality and frequency (Magnuson, Ruhm, & Waldfogel, 2007; Sackes et al., 2011; Skibbe et al., 2008). As the children age, they might begin to adopt increasingly disinterested or negative views toward science as the content becomes increasingly abstract and demanding as well as disconnected from their everyday experiences at home and school (Archer et al., 2010; Stanovich, 1986). Further, reading and mathematics achievement gaps experienced in the elementary grades might make it more difficult to acquire the scientific knowledge and skills concurrently being taught. Collectively, these early and ongoing experiences could result in children's following cumulative growth trajectories with the science achievement gaps gradually widening.

A third possibility is that children at risk follow a compensatory achievement growth trajectory (Huang, Moon, & Boren, 2014; Leppänen, Niemi, Aunola, & Nurmi, 2004). Here, and despite entering school with initially less informal knowledge about the physical and natural sciences, the children's science achievement gaps might decrease over time. The decreasing gaps might occur as children arrive at school and begin to receive instruction and support that help counter the fewer informal learning opportunities these children encountered prior to school entry. As a result, and despite an initial lag, achievement growth experienced by children at risk might begin to accelerate relative to their higher-achieving peers. This might especially occur for language-minority children as their English language proficiency increases (Crosnoe, 2012; Reardon & Galindo, 2009). Alternatively, these gains might be due to supplementary programs delivered to lower- but not to higher-achieving children (Skibbe et al., 2008) or a result of high-quality school climates that are especially beneficial to children in need of additional support, safety, and structure (Bryk, Sebring, Allensworth, Luppescu, & Easton, 2010). Thus, and although children following a compensatory model might initially display relatively lower science achievement at school entry, their science achievement gaps could gradually lessen.

However, whether and to what extent science achievement gaps display stable, cumulative, or compensatory growth trajectories are unknown. This has been due to the lack of large-scale, multivariate, and longitudinal studies, particularly those following diverse samples across the elementary and middle school grades. Establishing which types of science achievement growth trajectories are experienced by those who are at risk would help inform the timing of school-based science intervention efforts. For example, such interventions may need to be delivered early in children's school careers (e.g., by the primary grades) if science achievement gaps remain stable or follow a cumulative growth trajectory.

Modifiable Factors Hypothesized to Explain Science Achievement Gaps

Rigorously investigating whether children experience stable, cumulative, or compensatory science achievement growth trajectories should involve examining potentially modifiable factors that might explain the science achievement gaps associated with the children's race, ethnicity, language, and low-income status (Bali & Alvarez, 2004). Doing so should better identify potential targets of intervention. Theoretically, a number of factors relating to early constrained opportunities and propensities to learn about science may help explain science achievement gaps. These factors may limit children's initial knowledge, abilities, and interest in science, thereby impeding over-time growth in science achievement (e.g., Eccles, Wigfield, & Schiefele, 1998; Jones & Byrnes, 2006; Li et al., 2009; Muller et al., 2001). For example, entering school with comparatively less knowledge about the natural and social sciences, struggling with reading and mathematics, and having difficulty attending to classroom instruction or otherwise meeting classroom behavioral expectations (e.g., remaining task focused, organized, persistent) may result in lower achievement initially and throughout elementary and middle school (Duncan et al., 2007; Morgan et al., 2009; Morgan, Farkas, Tufis & Sperling, 2008; Reardon, 2011; Sackes et al., 2011), as may attending economically disadvantaged or racially segregated schools (Mickelson, Bottia, & Lambert, 2013; Von Secker & Lissitz, 1999). Being raised in a single-parent household might result in fewer educational materials at home, experiencing less adult supervision and lower educational expectations, and less frequent parent–child language interactions (Morsy & Rothstein, 2015). These might constrain children's initial science achievement as well as science achievement growth (Pong, Dronkers, & Hampden-Thompson, 2003). Young, struggling readers often acquire lower vocabularies (and in particular, less abstract, technical, or "academic" vocabularies often used during science instruction), have less general knowledge, display less cognitive ability (Cunningham & Stanovich, 1997), and are often less able to comprehend science texts and generate science-related inferences (Tate, Jones, Thorne-Wallington, & Hogrebe, 2012). In contrast, O'Reilly and McNamara (2007) found that being a skilled reader helped compensate for a lower level of background knowledge about science among older children. Separately, young children who struggle with mathematics may subsequently acquire less well-developed logical reasoning and problem-solving skills, resulting in more difficulties with scientific reasoning, thinking, and explanation skills (Batista & Matthews, 2002). Children with lower behavioral self-regulation (e.g., inattention to task) generally display lower academic achievement (e.g., Duncan et al., 2007). Thus, children's early knowledge about the natural and social sciences, self-regulatory behaviors, and reading and mathematics achievement may constitute modifiable factors that, if increased through school-based interventions, may help prevent or reduce the early onset of science achievement gaps.

Yet these interrelations have yet to be rigorously examined, particularly using multivariate longitudinal data and analyses (Byrnes & Miller, 2007). For example, and although extant work has found that reading and mathematics achievement relate to science achievement by middle and high school (Gustin & Corazza, 1994; Slykhuis & Park, 2006), and cross-sectional data indicate that reading and mathematics achievement correlate with science achievement in fifth grade (Maerten-Rivera et al., 2010), no longitudinal study has yet examined whether and to what extent greater reading and mathematics achievement may help explain greater science achievement over time. Engaging in self-regulatory behaviors may help children better attend to instruction and complete learning tasks (Blair, 2002; McClelland, Acock, & Morrison, 2006) and is known to uniquely predict greater reading and mathematics achievement in upper-elementary-age schoolchildren (e.g., Duncan et al., 2007). However, whether and to what extent these self-regulatory behaviors help explain science achievement gaps, particularly during the upper elementary and middle school grades, is presently unknown (Sackes et al., 2011). Limited English proficiency may interfere with comprehending classroom instruction and so limit initial science achievement. Yet dual-language learning might also result in greater science achievement. This is because, upon reaching a threshold of linguistic competence, bilingualism's greater cognitive demands are hypothesized to accelerate children's academic achievement growth (Ardasheva, Tretter, & Kinny, 2012; Kempert, Saalbach, & Hardy, 2011; Ricciardelli, 1992), resulting in compensatory achievement growth trajectories (Crosnoe, 2012; Reardon & Galindo, 2009).

Lower socioeconomic status (SES) may play a particularly important role in explaining science achievement gaps, including those disproportionately experienced by racial-, ethnic-, and language-minority children (McCoach et al., 2006; Reardon, 2011). Children growing up in low-SES families typically experience comparatively fewer early opportunities to learn about the natural and social sciences, in part because their parents often have lower educational levels and therefore less science knowledge themselves as well as fewer resources available to direct toward the children's cognitive and academic growth (e.g., Bradley & Corwyn, 2002; Hart & Risley, 1995; Sackes et al., 2011; A. Wang, Shen, & Byrnes, 2013). Racial, ethnic, and language minorities are far more likely to be raised in families experiencing economic disadvantage. For instance, children who are racial and ethnic minorities are about twice as likely to live in poverty as those who are White (U.S. Census Bureau, 2012b). Children raised in poverty often attend poorly resourced and racially segregated schools that further limit their academic opportunities (Anderman, 1998; Eccles et al, 1998; Li et al., 2009; Liu & Whitford, 2011; Ma & Wilkins, 2002; Von Secker, 2004). For example, Borman and Dowling (2010) reported that attending a school with high concentrations of economically disadvantaged or minority children predicted reduced achievement even after accounting for many factors, including prior achievement. Thus, directing interventions to lessen economic disadvantage and racial segregation, and doing so early in children's school careers, may constitute an additional way to prevent or reduce science achievement gaps in the United States. Family stress and investment factors (e.g., the quality of parent–child verbal interactions, the frequency of shared storybook reading, the parent's educational expectations for the child, the parent's involvement in the school) have been found to at least partially mediate

the relationship between economic disadvantage and lower achievement (e.g., J.-S. Lee & Bowen, 2006) as well as constitute modifiable factors that are responsive to preschool- and school-based interventions (Bali & Alvarez, 2004). School climate has also been theorized or reported to contribute to children's academic achievement (Bryk et al., 2010). Greater academic achievement should result when teachers and administrators are able to offer safe, supportive, and well-structured school environments that allow children to better enjoy and so benefit from their educational experiences (McCoy, Roy, & Sirkman, 2013; Stewart, 2008). Higher-quality school climates have repeatedly been observed to uniquely but modestly predict higher reading and mathematics achievement (e.g., Bodovski, Nahum-Shani, & Walsh, 2013; McCoy et al., 2013; Stewart, 2008) and may help explain achievement gaps (Fryer & Levitt, 2004), although findings about their relation with science achievement are inconsistent (Ma & Wilkins, 2002). In addition, school climate has often been studied as a time-invariant factor (e.g., Bodovski et al., 2013; Ma & Wilkins, 2002) in analyses without statistical control for other potential confounds (e.g., children's reading achievement or behavioral self-regulation), resulting in ambiguity as to its over-time and unique association with science achievement.

Study's Purpose and Research Questions

We investigated the early onset and over-time dynamics of science achievement gaps as well as potentially modifiable factors that may explain these gaps. To better address limitations in the field's current knowledge base, we analyzed a large longitudinal data set of children followed from kindergarten entry to the end of eighth grade. This data set included large samples of children from racial-, ethnic-, and language-minority groups and lower-SES families. These data also contain grade-appropriate science achievement scores for third-, fifth-, and eighth-grade children as well as general knowledge scores resulting from kindergarten and first-grade children's responses to both natural and social science achievement items. Accordingly, our investigation of the early onset and over-time dynamics of science achievement gaps includes the acquisition of general knowledge in kindergarten and first grade, followed by science achievement in third, fifth, and eighth grades. We investigated the following six specific research questions:

1. How large are general knowledge gaps occurring in kindergarten, and to what extent do these continue to occur by the end of first grade?
2. As children move from third to eighth grade, what is their typical initial level (i.e., intercept) and rate of achievement growth (i.e., slope) in science?
3. Are these gaps consistent with stable, cumulative (i.e., gap increasing), or compensatory (i.e., gap decreasing) achievement growth trajectories? How do these initial third-grade science achievement levels and third- to eighth-grade growth trajectories vary by children's race, ethnicity, language, and family SES status? How are a more general set of child- and family-level characteristics, including parenting quality, related to typical levels of third-grade science achievement in the United States as well as to achievement growth from third to eighth grade?

4. To what extent are the third-grade science achievement gaps, as well as third- to eighth-grade science achievement growth, explained by such modifiable factors as general knowledge, reading and mathematics achievement, and behavioral self-regulation? How much of children's later science achievement can be predicted by their first-grade achievement-related knowledge, skills, and behaviors?
5. With the aforementioned first-grade predictive factors accounted for, how important are the modifiable factors of children's subsequent reading and mathematics achievement, and behavioral self-regulation at each of third, fifth, and eighth grades to their science achievement during these grades?
6. To what extent does a school's academic climate and racial, ethnic, and economic composition explain children's science achievement, over and above the aforementioned child- and family-level factors?

Method

Study's Database

The database analyzed is the public-use file of the Early Childhood Longitudinal Study, Kindergarten Class of 1998–1999 (ECLS-K), a nationally representative cohort of children who entered kindergarten in 1998. (The sample was refreshed in first grade to maintain national representation.) The ECLS-K is maintained by the National Center for Educational Statistics (NCES), Institute of Education Sciences, U.S. Department of Education. This is the first large-scale, nationally representative, longitudinal sample of children followed at intervals as they matriculated through their elementary and middle school years. The sample was selected to be representative of all U.S. children entering kindergarten in the fall of 1998. Children were recruited from both public and private kindergartens offering full- or half-day classes. Data were collected in the fall of 1998, the spring of 1999, and the fall of 1999 (with data from only a smaller, random subsample collected at this time point) and again in the springs of 2000, 2002, 2004, and 2007. For most children, these time periods correspond to the fall and spring of kindergarten, the fall and spring of first grade, and the springs of third, fifth, and eighth grades. This study used data from the fall of kindergarten and the springs of first, third, fifth, and eighth grades. Table 1 displays the analytic sample's descriptive statistics.

Measures

Child and family characteristics. We included dummy variables in our analyses capturing a range of antecedent child and family sociodemographic characteristics. These included the child's race and ethnicity (non-Hispanic White, Black/African American, Hispanic, Asian, American Indian, and Other), gender, parents' marital status, and whether or not the primary language spoken in the home is English. The analyses also included a continuous, composite measure of SES, including the father's and mother's education levels (less than high school, high school graduate, some college, college graduate or more), occupations, and family

Table 1
Weighted Means and Standard Deviations or Percentages of Time-Invariant Variables (*N* = 7,757)

Variable	*M* or Percentage	*SD*
Child is White	57.77%	
Child is Black	16.57%	
Child is Hispanic	18.15%	
Child is Asian	2.89%	
Child is American Indian	1.60%	
Child is other race/ethnicity	3.01%	
Child is male	51.99%	
Child age in months at fall kindergarten	68.45	4.20
Child age in months at spring first grade	86.86	4.16
Family SES at kindergarten	0.012	0.78
Family SES at first grade	−0.036	0.79
Child's mother is not married at fall kindergarten	30.13%	
Child's mother is not married at spring first grade	29.47%	
Family speaks non-English at home at fall kindergarten	12.27%	
Reading Test score at fall kindergarten	35.54	9.95
Reading Test score at spring first grade	77.83	24.14
Mathematics Test score at fall kindergarten	26.24	8.96
Mathematics Test score at spring first grade	61.71	17.92
General Knowledge Test score at fall kindergarten[a]	22.76	7.56
General Knowledge Test score at spring first grade[b]	34.76	7.56
Approaches to Learning at fall kindergarten	3.00	0.66
Approaches to Learning at spring first grade	3.03	0.69

Note. SES = socioeconomic status.
[a]Sample size for this variable is 7,210.
[b]Sample size for this variable is 7,619.

income (less than $25,000; between $25,001 and $50,000; between $50,001 and $100,000; and greater than $100,000). We included child age at the data collection dates in fall of kindergarten and spring of first grade as continuous variables.

School demographic characteristics. We used the percentage of minorities to represent the school's racial composition and the percentage of free-lunch-eligible students to indicate the school's overall SES of those enrolled.

School academic climate. We used surveys of teachers and administrators to create school academic climate variables. These variables were summed at the respective grade levels and *z* scored. During the third, fifth, and eighth grades, school administrators were asked about the extent to which they placed emphasis on (a) achieving high standards, (b) curricula that aligned with high standards, (c) instructional strategies that aligned with high standards, and (d) teacher professional development. The values for these variables ranged from 1 = *no/minor emphasis* to 3 = *major emphasis*.

During the third, fifth, and eighth grades, school administrators were asked if they thought there were problems in the school's neighborhood with gangs, drug use, vacant buildings, and crime. The values for these variables ranged from 1 = *big problem* to 3 = *no problem*. School administrators were also asked if they had a problem with teacher turnover.

During these same survey waves, teachers were asked if they thought misbehavior in the classroom affected their teaching and whether their students are incapable of learning. The values for the variables ranged from 1 = *strongly disagree* to 5 = *strongly agree*. Teachers were also asked if they felt fellow staff members accepted them as a colleague, if the staff had school spirit, whether they enjoyed their present teaching job, and if they were satisfied with their class size. The values for the variables ranged from 1 = *strongly disagree* to 5 = *strongly agree*. Teachers were asked their highest level of educational achievement. The values for the variables ranged from 1 = BA/BS to 4 = ED/PhD. Teachers were asked how long they have been teaching. This was a continuous variable that ranged from 0 to 35 years. The third-grade variables were summed and then *z* scored to create the value of the third-grade time-varying school academic climate variable. The fifth- and eighth-grade variables were summed and *z* scored to create the time-varying fifth- and eighth-grade school academic climate variables, respectively.

General knowledge. A General Knowledge Test was used to assess children's understanding of the physical, biological, and social sciences. This measure was administered in the fall and spring semesters of kindergarten and the spring semester of first grade. The General Knowledge Test was the ECLS-K's best measure of children's initial knowledge about science (Sackes et al., 2011). It included equal proportions of questions on (a) the earth, physical, and life sciences and (b) social studies. (Unfortunately, NCES did not make subject-specific subscores available.) The science-related items focused on two types of competencies: (a) conceptual understanding of earth and space science, life science, and physical science concepts and (b) scientific process skills, such as asking questions, deriving conclusions, and making predictions. We found that children's first-grade General Knowledge Test scores were highly predictive of their third-grade Science Test score, helping to justify use of the kindergarten and first-grade General Knowledge Tests as measures of general science achievement. The General Knowledge Test was field tested, and the psychometric properties of items were evaluated using item response theory (IRT) methods. NCES used these IRT methods to create adaptive tests that were individually administered in an untimed format. Children were given a test with coverage varying according to their grade and skill level (Rock & Pollack, 2002). Reliability of the IRT scale scores obtained in fall of kindergarten and spring of first grade were .88 and .89, respectively (Najarian, Pollack, & Sorongon, 2009).

Science achievement. The Science Test was administered in third, fifth, and eighth grades. The test specifications for these measures were developed based on the Science Framework of the 1996 National Assessment of Educational Progress (NAEP; Allen, Carlson, & Zelenak, 1999). This included two broad

classes of science competencies: (a) conceptual understanding, which refers to both factual knowledge and conceptual descriptions of why things occur as they do, and (b) scientific investigation, which refers to the ability to form questions about the natural world, use basic tools and available evidence to answer those questions, and communicate the answers and the processes used. At each of these three survey waves, questions were drawn in approximately equal proportions (i.e., about 33%) from the fields of earth, physical, and life science (Najarian et al., 2009). As with the ECLS-K's other academic achievement tests (e.g., the Reading Test, the Mathematics Test), each of the Science Test's items was field tested and included only if it displayed acceptable properties. Specifically, items were included in the final form of the test if they displayed (a) acceptable item-level statistics, (b) good fit with maximum-likelihood IRT, and (c) no differential item functioning across gender or race. Reliabilities of the IRT scale scores for spring third, fifth, and eighth grades were .88, .87, and .84, respectively (Najarian et al., 2009).

Reading achievement. The ECLS-K Reading Test measures reading proficiencies appropriate for the grade level of assessment. The Reading Test was created through a multistage panel review. Some items were borrowed or adapted from published tests (e.g., the Peabody Picture Vocabulary Test–Revised, the Woodcock Johnson Tests of Achievement–Revised). The Educational Testing Service, elementary school curriculum specialists, and practicing teachers supplied other items. All items were field tested and psychometric properties of items were evaluated using IRT methods. During kindergarten and first grade, items measured basic skills (print familiarity, letter recognition, beginning and ending sounds, rhyming sounds, word recognition, receptive vocabulary) and comprehension (listening comprehension, words in context). During the third, fifth, and eighth grades, items measured skill in initial understanding and interpretation, personal reflection and response, and demonstrating a critical stance. The reliabilities of the IRT scale scores for the kindergarten through eighth-grade assessments ranged from .87 to .96 (Najarian et al., 2009). Construct validity was supported by a high correlation (i.e., .83 for spring of third grade) between Reading Test scores and children's scores from the Woodcock-McGrew-Werder Mini-Battery of Achievement (MBA; Woodcock, McGrew, & Werder, 1994).

Mathematics achievement. The Mathematics Test was also administered during each of the ECLS-K survey waves. NCES used a multistep panel review process to develop a Mathematics Test for use in the ECLS-K. This test was based on the NAEP's specifications. The kindergarten and first-grade items included counting, identifying numbers and shapes, place value, measurement, sequencing, and basic operations. The third- and fifth-grade items included measurement, number sense, properties, and operations. More advanced items in eighth grade included geometry, spatial sense, data analysis, statistics, probability, patterns, algebra, and functions. A wide range of test bank items was used. Criterion-referenced clusters of items were associated with specific stages that children passed through from premathematics to mathematics achievement. Reliabilities of the IRT-scaled

scores were in the mid-.90s for each of the assessment waves (Najarian et al., 2009). The correlation between the Mathematics Test scores and children's scores from the MBA (Woodcock et al., 1994) was .84 for spring of third grade, indicating good construct validity.

Approaches to Learning. Teachers completed a modified version of the Social Skills Rating System (SSRS; Gresham & Elliott, 1990) during the kindergarten and first-, third-, and fifth-grade ECLS-K survey waves to rate children's behavior. The original psychometric data for the SSRS indicated a .85 test–retest correlation across 4 weeks (Gresham & Elliott, 1990). Construct validity was supported by both correlational and factor analyses (Feng & Cartledge, 1996; Furlong & Karno, 1995). NCES subsequently modified the SSRS; this modified form is known as the Social Rating Scale. General education teachers in kindergarten and first, third, and fifth grade used a frequency scale to rate how often the child displayed a particular social skill or behavior (i.e., 1 = *never* to 4 = *very often*). Items used for the Approaches to Learning subscale measured how well a child self-regulated his or her behavior while completing learning-related tasks (e.g., remaining attentive, persisting at task, being flexible and organized). Because these data were not collected from eighth-grade teachers, we imputed each child's fifth-grade score for eighth grade. This should have provided the best available eighth-grade measure of this variable. The Approaches to Learning subscale is the ECLS-K's best measure of learning-related or self-regulatory behavioral functioning. Controlling for prior achievement, these behaviors best predict children's later achievement (Duncan et al., 2007; Tach & Farkas, 2006). Pollack, Atkins-Burnett, Najarian, and Rock (2005) report split-half reliabilities for the Approaches to Learning subscale of .77 to .89.

Parenting quality. In the third-grade survey wave, parents were asked how often they told stories, sang songs, helped with art, gave chores to do, played games, taught about nature, built things, did sports, practiced numbers, and read stories with their child. Values of these 10 variables were converted into a per-week value, summed, and z scored. Parents were also asked if they engaged in activities with their child or their child engaged in activities, such as attending a play, visiting a museum, visiting a zoo, visiting a library, attending a sporting event, being taken to music or art lessons, and performing in plays. Parents were also asked if the family received newspapers and magazines in the home. Other questions asked if the family had a dictionary, owned a calculator, had a library card, had a library card for the child, had a special place for homework, and had someone to help the child with homework. The values of these 18 variables were summed (with 1 = *yes*, 0 = *no*, for each variable) and z scored.

Parents were asked how many children's books were in the home. The maximum number of books was chosen as 1,000 as there were a few high values present. The number of books was z scored. Parents were asked how often their child did homework and the time set aside for homework every day. Each of these two variables was z scored, summed, and then z scored again. Parents were asked how far they expected their child would go with regard to their education. This variable was z scored. Parents

were also asked if they were involved in school activities, such as attending an open house, attending a parent–teacher conference, volunteering at the school, helping with fundraising, or attending a school event. The values of these five variables were summed (with 1 = *yes*, 0 = *no*, for each variable) and *z* scored. These parenting items were then summed, *z* scored again, and used as a parenting quality variable.

Analyses

Missing data. Growth modeling allows a child to remain in the analytical sample and contribute to coefficient estimates so long as he or she has usable data on the dependent variable for at least one time period. We used IVEWARE, an SAS callable program to multiply impute missing values of variables used to predict our dependent variables. Five data sets were created in the imputation process, and results for each data set were combined (Raghunathan, Lepkowski, Van Hoewyk, & Solenberger, 2001).

Analytic strategies. We began the data analysis by regressing fall kindergarten General Knowledge Test scores against SES, race and ethnicity, home language use, and other sociodemographic background variables. This indicated SES and race differentials in General Knowledge Test scores at kindergarten entry. Next, we regressed spring first grade general knowledge scores, first against sociodemographics and then after adding fall kindergarten general knowledge, mathematics, and reading achievement as well as behavioral self-regulation to the equation. This showed the magnitudes of the disparities in general knowledge at the end of first grade as well as the extent to which they are accounted for by general knowledge gaps at kindergarten entry. These results are reported in Table 2, and they answer Research Question 1 concerning general knowledge gaps in kindergarten and first grade.

Next, we computed average values for third-, fifth-, and eighth-grade science achievement test scores and the other time-varying variables. Science achievement test score averages are reported in Table 3 and are then plotted separately for race-ethnicity groups (Figure 1), SES quintiles (Figure 2), and whether or not English is spoken at home (Figure 3). This table and these figures provide preliminary answers to Research Questions 2 and 3 about science achievement growth trajectories and how they differ across sociodemographic groups.

Following this, we used multilevel growth modeling (Raudenbush & Bryk 2002; Singer & Willett, 2003) to identify factors associated with or predictive of science achievement growth trajectories across third, fifth, and eighth grades and to relate first-grade general knowledge to subsequent science achievement. For these models, Level 1 contained the repeated science achievement scores for each child, Level 2 contained unchanging child-level variables, and Level 3 included the kindergarten classrooms in which these children were originally clustered.

Three growth models were estimated and are reported in Table 4. The first model includes only unchanging sociodemographic variables as predictors. This answers most of Research Question 3 about the specifics of science growth trajectories and how they differ across sociodemographic groups. By including Time and Time² in the equation, we allowed for nonlinear growth patterns.

The second growth model reported in Table 4 added first-grade measures of general knowledge, reading and mathematics achievement, and behavioral self-regulation to the equation. This answered Research Question 4 about the ability of prior academic performance and behaviors to predict subsequent science achievement.

The final model in Table 4 added time-varying reading and mathematics achievement, behavioral self-regulation, percentages minority and free lunch in the school, and school academic climate as well as a third-grade measure of parenting quality to the equation. This answered the remaining parts of Research Question 3 as well as Research Questions 5 and 6. We present this model in detail, since the other two simpler models are nested within it.

This most complex of the estimated growth models involves a Level 1 model that has Time and Time² on the right side of the equation as well as the time-varying variables of reading achievement, mathematics achievement, behavioral self-regulation, and two time-varying school composition variables (i.e., percentage minority and percentage free-lunch-eligible students). (Note that we treat the latter two variables as child level and time varying because the children have moved to many different schools and few are clustered in the same school by these grades.) Thus, the Level 1 equation is

$$\text{Science Test Score}_{tij} = a_{0ij} + a_{1ij}\text{Time}_{ti} + a_{2ij}\text{Time}_{tij}^2 + a_{3ij}\text{Reading Test Score}_{tij} + a_{4ij}\text{Mathematics Test Score}_{tij} + a_{5ij}\text{Approaches to Learning}_{tij} + a_{6ij}\text{Percentage Minority in School}_{tij} + a_{7ij}\text{Percentage Free Lunch in School}_{tij} + e_{tij}, \quad (1)$$

where *t* refers to the time, and *i* and *j* indicate the *i*th child in the *j*th kindergarten. Here, person-specific variables at third, fifth, and eighth grades predicted person-specific Science Test scores at each of these time periods. The error term was independently normally distributed, with a mean of zero. Then, the Level 2 equations showed the intercept and slope (against time) coefficients as varying and dependent on unchanging person-specific sociodemographic and first-grade reading and mathematics achievement and behavioral self-regulation variables, with the other Level 1 coefficients taken as fixed:

$$a_{0ij} = b_{0j} + b_{1j}\text{Sociodemographics}_{ij} + b_{2j}\text{Test Scores / Approaches Spring First Grade}_{ij} + r_{0ij}, \quad (2)$$

$$a_{1ij} = c_{0j} + c_{1j}\text{Sociodemographics}_{ij} + c_{2j}\text{Test Scores / Approaches Spring First Grade}_{ij} + r_{1ij}, \quad (3)$$

where a_{2ij} and the other *a* coefficients with subscripts greater than 2 are fixed across individuals.

In these Level 2 models, we took the variance/covariance matrix of the residuals to be unstructured so that it was estimated from the data with no constraints. Finally, the Level 3 model allowed b_{0j} and c_{0j} to vary across the kindergartens that the children were originally clustered into:

$$b_{0j} = g_0 + u_{0j} \text{ and } c_{0j} = h_0 + v_{0j} \text{ with the other Level 2 coefficients fixed.} \quad (4)$$

Table 2
Standardized Parameter Estimates Predicting General Knowledge Test Scores, Weighted ($n = 7{,}210$)

Variable	Fall Kindergarten General Knowledge Score	Spring First-Grade General Knowledge Score	
		Model 1	Model 2
Intercept	.13***	.19***	.09***
Child is Black	−.62***	−.65***	−.28***
Child is Hispanic	−.29***	−.29***	−.06**
Child is Asian	−.41***	−.32***	−.11*
Child is American Indian	−.52***	−.55***	−.20***
Child is other race/ethnicity	−.06	−.06	−.05
Family SES at kindergarten	.39***	.36***	.06***
Child age in months at fall kindergarten	.28	.21***	.00
Child is male	−.01***	.07***	.12***
Child's mother is not married at fall kindergarten	−.12***	−.10***	−.00
Family speaks non-English at home at kindergarten	−.34***	−.61***	−.03
General Knowledge Test at fall kindergarten			.58***
Reading Test at fall kindergarten			.08***
Mathematics Test at fall kindergarten			.03***
Approaches to Learning, fall kindergarten			.05***
R^2	.38	.38	.64

Note. SES = socioeconomic status.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 3
Weighted Means or Percentages and Standard Deviations of Time-Varying Variables ($n = 7{,}731$–$7{,}757$)

Variable	M or Percentage	SD
Third grade		
Science Test score ($n = 7{,}731$)	50.53	15.28
Reading Test score	126.71	28.88
Mathematics Test score	99.06	24.90
Free-lunch percentage in school	37.44%	27.41
Minority percentage in school	45.72%	30.71
Approaches to Learning	3.01	0.66
Fifth grade		
Science Test score ($n = 7{,}739$)	64.37	16.02
Reading Test score	149.45	27.13
Mathematics Test score	122.92	25.38
Free-lunch percentage in school	38.82%	27.45
Minority percentage in school	47.14%	30.61
Approaches to Learning	3.03	0.68
Eighth grade		
Science Test score ($n = 7{,}741$)	83.08	16.76
Reading Test score	167.36	28.92
Mathematics Test score	139.45	23.28
Free-lunch percentage in school	35.00%	24.57
Minority percentage in school	48.49%	29.39
Approaches to Learning	3.03	0.68

This multilevel model was the most complex one we estimated. We built up to this model by first estimating a less complex model in which the only predictors are unchanging sociodemographic variables. The estimated effects from this less complex model showed the total (i.e., gross) differences on Science Test score intercepts and slopes between groups defined by social class, race and ethnicity, and the other sociodemographic variables. Following this, we added the spring of first-grade General Knowledge, Reading, and Mathematics Test scores as well as the Approaches to Learning score to the model. This showed the relative importance of these variables as predictors of science achievement growth trajectory intercepts and slopes as well as the extent to which these variables accounted for the sociodemographic group differences found in the first model. That is, sociodemographic differences in this model were adjusted for children's spring of first-grade general knowledge and reading and mathematics achievement as well as their behavioral self-regulation. As we shall see, these calculations suggested a strong role of first-grade general knowledge achievement gaps in predicting third-grade science achievement gaps.

Finally, we added the third-grade parenting and time-varying reading and mathematics achievement, behavioral self-regulation, and school characteristics (i.e., academic climate, percentage free lunch, and percentage minority) to the equation. This showed the strength of association between these variables and science achievement as well as the extent to which these variables accounted for remaining sociodemographic group differences in U.S. schoolchildren's science achievement. Sociodemographic
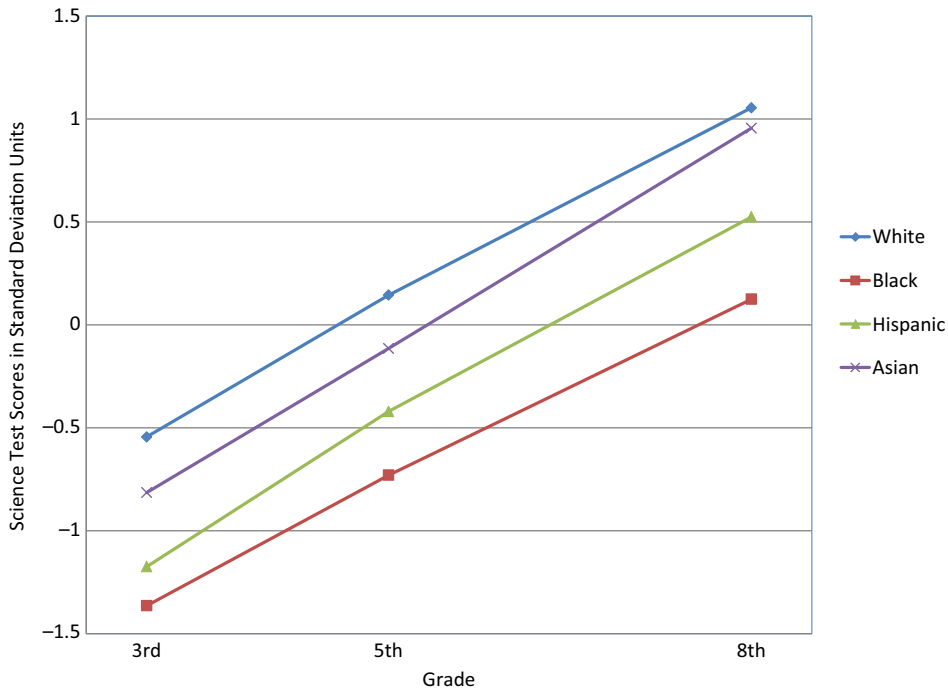
FIGURE 1. *Science Test score trajectories by racial and ethnic group*
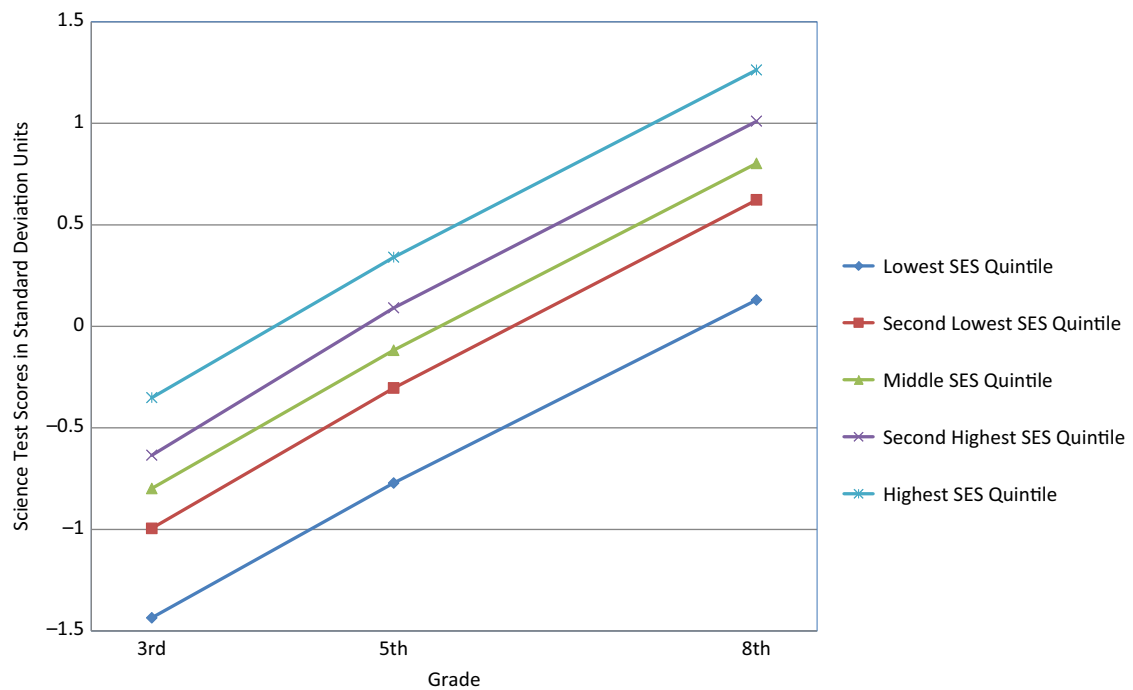

FIGURE 2. *Science Test score trajectories by socioeconomic status (SES) quintile groups (SES measured in first grade)*

differences in this model were adjusted both for the first-grade achievement and behavior models and for the time-varying achievement, behavior, and school-level variables. We selected these models because they contained a sufficient level of complexity to address our research questions without increasing this complexity beyond what was feasible with the available data. In particular, we followed the advice of Singer and Willett (2003,

p. 169) that when estimating models containing time-varying variables with only three waves of data, it is appropriate to estimate growth models in which individual intercepts and slopes are random and conditional on prior unchanging variables, but it is better to specify the coefficients of the time-varying covariates as fixed rather than random. All analyses were performed with PROC MIXED in SAS 9.3. We incorporated sampling weights
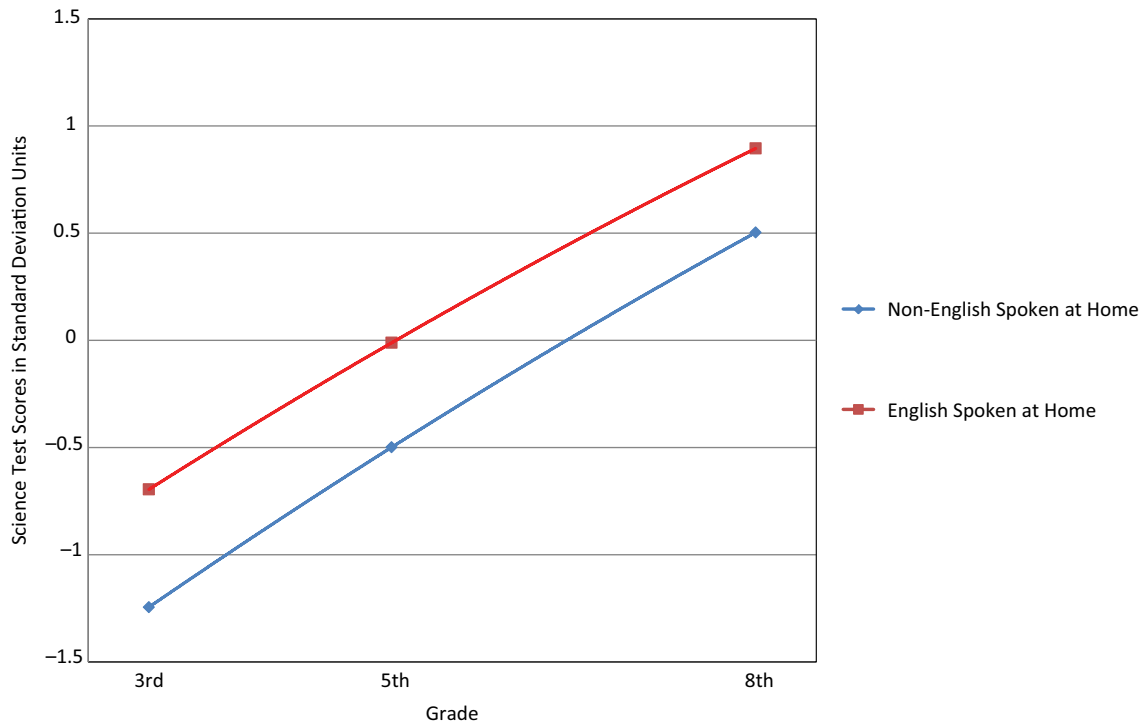
FIGURE 3. *Science Test score trajectories by English use in home (measured in fall of kindergarten)*

and design effects to account for oversampling of some population subgroups and for the stratified cluster design of the ECLS-K.

Results

Table 1 displays descriptive statistics, weighted to be nationally representative, for the fall kindergarten and spring first-grade sociodemographic, achievement, and behavior variables in our analytic sample of 7,757 students who had at least one Science Test score. The race and gender composition of this analytic sample were within 4 percentage points, and the mean SES score was within .05 standard deviations, of those of the full ECLS-K sample. (Table 1 omits the composite parenting variable as it was constructed as a *z* score.)

Table 2 displays the results of two sets of regressions. The first used sociodemographics to predict general knowledge in the fall of kindergarten. The second repeated this calculation to predict these scores for spring of first grade and then added fall kindergarten academic achievement and teacher-judged self-regulatory behaviors to the equation. The results of these analyses indicated, first, the relations between sociodemographics and General Knowledge Test scores when children entered kindergarten, including the magnitudes of general knowledge differentials (i.e., gaps) across race, ethnicity, language use, and SES groups. Second, they indicated how the magnitudes of the racial, ethnic, language use, and SES gaps changed by the spring of first grade. Finally, they showed the relation between (a) academic achievement and self-regulatory behaviors at kindergarten entry and (b) general knowledge in the spring of first grade. In particular, they showed the extent to which fall kindergarten academic achievement and behaviors accounted for general knowledge gaps in the spring of first grade.

The first column of Table 2 indicates large and significant general knowledge gaps at kindergarten entry. Children who are Black scored .62 of a standard deviation lower than children who are White. The gap was .29 of a standard deviation for Hispanics. Racial and ethnic gaps were evident despite statistical control for SES and for non-English language use in the home, which themselves were associated with, respectively, higher and lower science achievement. For Americans Indians, the gap was .52 of a standard deviation, and for Asians, it was .41 of a standard deviation. The coefficient for the continuous measure of SES was .39. That is, a 1–standard deviation change in SES was associated with a .39–standard deviation change in general knowledge. Thus, for example, Black children from families 1 standard deviation below the average SES entered kindergarten on average 1.01 standard deviations below the general knowledge score for White children whose families were at the SES mean. This is a large knowledge gap for children to face when they enter kindergarten.

The second column of this table repeats these analyses for general knowledge scores in the spring of first grade. At this grade level, the Black–White general knowledge gap increased to −.65 of a standard deviation. The Hispanic–White gap was constant at −.29 of a standard deviation, the gap for American Indians was now −.55 of a standard deviation, and the Asian–White gap was now −.32 of a standard deviation. The third column shows that large proportions of these first-grade gaps are accounted for by kindergarten entry academic achievement and behaviors. Among these factors, general knowledge has by far the strongest effect, with a standardized coefficient (effect size) of .58, whereas the coefficients of the other variables are all .08 or smaller. When these other variables were added to the regression equation, the gap for Black children decreased by 57%, the gap

Table 4
Table 4
Weighted Parameter Estimates of Three-Level Growth Models Predicting Science Achievement
(n = 7,731–7,757)

Variable	Model 1	Model 2	Model 3	% Reduction
Intercept	−.72 ***	−.80 ***	−.53 ***	26.39
Time	.03 ***	.03 ***	.01 ***	66.67
Time × Time	−.0001 ***	−.0001 ***	.0001 ***	200.00
Time-varying Reading Test			.24 ***	
Time-varying Mathematics Test			.26 ***	
Time-varying Approaches to Learning rating			.014	
Time-varying percentage minority in school			−.02 ***	
Time-varying percentage of free lunch in school			−.005	
Time-varying school academic climate			.001	
Effects on intercept				
Child is Black	−.52 ***	−.19 ***	−.13 ***	75.00
Child is Hispanic	−.23 ***	−.06 ***	−.04 **	81.30
Child is Asian	−.08 *	.01	.002	102.50
Child is American Indian	−.45 ***	−.14 ***	−.03	93.33
Child is other race/ethnicity	−.13 **	−.05 *	−.04	69.23
Family SES at first grade	.30 ***	.05 ***	.01	96.67
Child age in months at spring first grade	.08 ***	−.02 ***	.001	98.75
Child is male	.16 ***	.15 ***	.12 ***	25.00
Child's mother is not married at spring first grade	−.06 ***	−.02	−.01	83.33
Family speaks non-English at home at fall kindergarten	−.24 ***	.10 ***	−.06 ***	75.00
General Knowledge Test at spring first grade		.42 ***	.31 ***	
Reading Test at spring first grade		.09 ***	−.01	
Mathematics Test at spring first grade		.09 ***	−.06 ***	
Approaches to Learning, spring first grade		.02 ***	−.03 ***	
Parenting quality, third grade			.001 ***	
Effects on slope				
Child Is Black × Time	−.002 ***	−.002 ***	−.001 ***	50.00
Child Is Hispanic × Time	.001 **	.001	.0005	50.00
Child Is Asian × Time	.002 ***	.002 **	.001	50.00
Child Is American Indian × Time	.0003	−.0003	−.001	333.33
Child Is Other Race/Ethnicity × Time	−.001	−.001	−.001	0.00
Family SES at First Grade × Time	.0004 *	.001 ***	.0005 ***	−25.00
Child Age in Months at Spring First Grade × Time	−.001 ***	−.001 ***	−.001 ***	0.00
Child Is Male × Time	−.0005 *	−.0002	.0005 **	200.00
Child's Mother Is Not Married at Spring First Grade × Time	−.0005	−.0005 *	−.001 ***	80.00
Family Speaks Non-English at Home at Fall Kindergarten × Time	.001 **	.0001	−.0002	120.00
General Knowledge Test at Spring First Grade × Time		−.002 ***	−.002 ***	
Reading Test at Spring First Grade × Time		−.0002	.0005 ***	
Mathematics Test at Spring First Grade × Time		.001 ***	.002 ***	
Approaches to Learning at Spring First Grade × Time		.001 ***	.001 ***	
Parenting Quality, Third Grade × Time			−.00004 ***	
AIC	36,338.8	29,317.5	26,665.8	
BIC	36,496.6	29,526.0	26,568.0	

Note. All continuous independent variables have been standardized. Time-varying dependent and independent variables were standardized using the means and standard deviations of the data pooled across third, fifth, and eighth grades. Time coded in months, with 0 = January 1, 2002. Percentage reduction is from Model 1 to Model 3. Statistically significant findings carried out to nearest nonzero decimal point to show effect size. SES = socioeconomic status; AIC = Akaike information criterion; BIC = Bayesian information criterion.
*p = .05. **p = .01. ***p = .001.

for Hispanics decreased by almost 80%, and that for American Indians decreased by about 64%. The gap for speaking a language other than English at home decreased by almost 100% and was fully explained. With these additional variables controlled, the gap associated with SES decreased by 83%. As indicated by the values of R^2, 64% of the variance in children's spring of first-grade general knowledge was explained by Model 2's factors. Thus, the general knowledge gaps initially observed

on the basis of children's race, ethnicity, language use, and family SES in the spring of first grade were mostly attributable to the large size of these gaps already occurring at kindergarten entry.

We observed a strong relation between children's general knowledge in the fall of kindergarten and the spring of first grade. In third, fifth, and eighth grades, NCES administered a measure consisting of items exclusively designed to assess children's science achievement. We applied multilevel modeling to these data to estimate science achievement growth curve trajectories from third to eighth grade. We used two kinds of variables as predictors in these models. The first of these were sociodemographics, general knowledge, reading and mathematics achievement, behavioral self-regulation, and parenting measured during spring of first grade, which served as unchanging predictors. The second group of variables included time-varying reading and mathematics achievement, behavioral self-regulation, school demographics, and academic climate characteristics, which served as time-varying covariates.

Table 3 shows the weighted means or percentages of the time-varying variables. The scale scores for the measures of science, reading, and mathematics achievement increased in a regular manner, while the school composition and self-regulatory behavioral variables were relatively constant over time. (We omitted the means for the academic climate variable because we constructed it as a z score.)

Figures 1, 2, and 3 display science achievement growth trajectories, in standard deviation units (based on the mean and standard deviation of all Science Test scores, pooled across third, fifth, and eighth grades), separately for the racial, ethnic, family SES, and language-use groups. (Data plotted are the means of Science Test scores, in standard deviation units, for each group for each grade.) The stability of these trajectories was striking, indicating approximately parallel growth trajectories throughout elementary and middle school. That is, the groups (whether by race, ethnicity, SES quintiles, or language use) differed more in their starting achievement levels in third grade than in their subsequent growth rates up through eighth grade. Comparing the science achievement of the minority racial and ethnic groups to that of Whites, non-English- to English-speaking households, and the lower to that of the higher SES quintile yields a consistent conclusion that children's initial science achievement was highly predictive of their later science achievement. Exceptions to this pattern were that, by eighth grade, Asian children largely closed, and Hispanic children modestly narrowed, the gap with White children. In contrast, the Black–White gap modestly increased across these grades. Among the SES groups, the top four quintiles maintained a relatively high degree of parallel growth, but the trajectory of the lowest SES quintile is an exception, showing not only a particularly large gap in third grade but also a modest widening of this gap by eighth grade. Overall, the science achievement gaps between White and Asian children and, to a lesser extent, between White and Hispanic children and children from non-English versus English-speaking homes were generally consistent with a compensatory model of achievement growth. In contrast, the science achievement gaps between Black and White children, and between low- and high-SES children, displayed trajectories more consistent with a cumulative model of achievement growth.

To identify the mechanisms responsible for these growth curves, Table 4 shows the results of fitting a series of increasingly complex three-level growth models in which both the intercepts and slopes of these curves were expressed as a function of both unchanging sociodemographics, parenting, and academic and behavioral functioning measured in first grade. The most complex of these models also included the following time-varying covariates measured in third, fifth, and eighth grades: children's reading and mathematics achievement; their behavioral self-regulation; the racial, ethnic, and economic composition of their schools; and a measure of the quality of the school academic climate. In these analyses, the continuous variables were made into z scores. In the case of time-varying variables (including the Science Test score dependent variable), this was done by pooling observations across third, fifth, and eighth grades, computing the means and standard deviations of the pooled data, and using these values to create z scores for these variables. As a consequence, all regression coefficients were standardized and can be directly compared to one another and treated as effect sizes. The large declines in the Akaike and Bayesian information criteria as we move from Model 1 to Model 3 of this table shows that the later models greatly improved the fit to the data.

Model 1 of Table 4 indicated that in standardized terms, the intercept for Science Test scores (the value on January 1 of the third-grade school year, when $t = 0$ in the model) is $-.72$ of a standard deviation (note that Science Test scores have been standardized; their overall mean across third, fifth, and eighth grades is zero). The growth rate in this model is .03 of a standard deviation per month, or .36 of a standard deviation per year. The negative coefficient on Time^2 shows that the growth rate decreased over time, but the magnitude of this coefficient indicated a very small effect.

Examining the coefficients relating the sociodemographic variables to the intercept, we see that on January 1 of third grade, and controlling for the other variables, the gaps for the racial- and ethnic-minority groups as compared to Whites were as follows, in standard deviation units: Blacks, $-.52$; Hispanics, $-.23$; Asians, $-.08$; and Native Americans, $-.45$. For every 1–standard deviation increase in family SES, the average child's Science Test score increased by .30 of a standard deviation. The third-grade science achievement of males was .16 of a standard deviation higher than that to females, and those of children with unmarried mothers were .06 of a standard deviation lower compared to those with married mothers. The science achievement of children from non-English-speaking homes was .24 of a standard deviation lower on average compared to those from English-speaking homes.

The relation between sociodemographics and the science achievement growth rate (slope) were of much smaller magnitude. Children who are Black displayed significantly lower growth rates than children who are White, but the difference is small in magnitude. By contrast, the science achievement of Hispanics, Asians, and those from non-English-speaking homes grew slightly faster than that of Whites or those from English-speaking homes. To see how these estimates translate into science achievement gaps by the spring of eighth grade, consider that on average, Black children were .52 of a standard deviation below White children on January 1 of third grade and 5 years

later are an additional 60 months × −.002 = .12 of a standard deviation lower than White children. Thus, most of the eighth-grade Black–White science achievement gap was already evident in third grade (where, as we have seen, it was strongly predicted by the first-grade general knowledge gap). Undertaking similar calculations, we found that the intercept for Hispanic children was .23 of a standard deviation lower than for White children, but by spring of eighth grade, they narrowed this gap by 60 months × .001 = .06 of a standard deviation. Asian children began .08 of a standard deviation lower than White children, but by eighth grade, this gap had been reduced by 60 months × .002 = .12 of a standard deviation. Children from non-English-speaking homes displayed .24 of a standard deviation lower science achievement than children from English-speaking homes by third grade, but by eighth grade, this gap had been reduced by .06 of a standard deviation. An increase of 1 standard deviation in family SES was associated with a science achievement increase in third grade (relative to the overall mean) of .30 of a standard deviation, whereas eighth-grade achievement increased relative to the overall mean by only .0004 × 60 = .02 of a standard deviation. Thus, for language use and SES, as for race and ethnicity, much of the science achievement gap between lower- and higher-achieving eighth-grade children was already in place by third grade.

Model 2 added first-grade general knowledge, reading, mathematics, and self-regulatory behavioral functioning to the equation. All four variables were statistically significant, but by far the largest effect on the third-grade intercept, .42, is for general knowledge. Thus, when first-grade general knowledge increased by 1 standard deviation, the associated increase in third-grade science achievement was .42 of a standard deviation. By contrast, the predicted effects for both reading and mathematics achievement were less than 25% of this magnitude, and the .02 for behavioral self-regulation was particularly small. With these variables controlled, much of the sociodemographic science achievement gaps was explained. In particular, 83% of the SES gap, 58% of the language-use gap, 63% of the Black–White gap, 73% of the Hispanic–White gap, and 69% of the American Indian–White gap were explained by these first-grade variables. As with Model 1, Model 2 shows that these variables had much smaller effects on the over-time growth rate of science achievement. That is, after entering these controls, most of the relation between children's sociodemographics and their science achievement was already evident by first grade.

Model 3 added the time-varying variables to the equation (i.e., children's reading achievement, mathematics achievement, and behavioral self-regulation along with the racial, ethnic, and economic composition as well as the academic climate of their schools). This model also added the unchanging parenting variable, which was measured in third grade. Over-time reading and mathematics achievement was strongly associated with over-time science achievement, with predicted effect sizes of approximately .25. By contrast, the effects of time-varying self-regulatory behaviors and school-level variables were small or nonsignificant. With these time-varying variables in the equation, the magnitude of the average science achievement growth rate declined by almost two thirds, from .03 of a standard deviation per month to .01 of a standard deviation per month. That is, two thirds of the science

achievement growth rate was attributable to reading and mathematics achievement growth. (Percentage reductions from Model 1 to Model 3 are shown in the final column of the table.) Further, with the time-varying variables controlled, the already small sociodemographic differences in science growth rates became even smaller. The third-grade measure of parenting showed a positive and significant effect on the intercept (i.e., third-grade Science Test score), and with this variable controlled, the effects of race and ethnicity and SES were further explained.

To summarize, large gaps in general knowledge were already evident when children entered kindergarten and remained evident in the spring of first grade. These factors strongly predicted the large science achievement gaps in third grade, and these gaps persisted through eighth grade. Persistent science achievement gaps from third through eighth grades were attributable to lower reading and mathematics achievement as children moved through third, fifth, and eighth grades. Accounting for these and the study's many other measured factors, including parenting, fully or substantially explained the large science achievement gaps initially observed for children at risk (e.g., racial-, ethnic-, and language minorities; those from lower-SES families) by eighth grade in the United States.

Discussion

To date, science achievement gaps in the United States have been frequently observed but rarely examined over time or explained. In particular, the age of onset and over-time stability of these gaps, as well as whether and to what extent they may be explained by potentially modifiable factors (e.g., lower general knowledge, lower reading and mathematics achievement, inattention and other self-regulatory behavioral difficulties, attending an economically disadvantaged or racially segregated school) has been largely unknown (Byrnes & Miller, 2007). This lack of a knowledge base has constrained efforts to address the nation's science achievement gaps. Yet doing so is critically important. Low levels of science achievement are no longer a "gathering storm" but now are "rapidly approaching a Category 5" (NASNAEIM, 2010, p. 1) in their potential to derail the nation's long-term global competitiveness (NASNAEIM, 2010, 2011). If left unaddressed, and given the nation's increasing economic disparities, low science achievement may be experienced by growing segments of the U.S. adult population. The result may be an electorate with a more limited ability to understand pressing public policy issues necessitating increasingly greater scientific literacy as well as lower employment and economic prosperity (NASNAEIM, 2010, 2011).

Collectively, results from our study indicate that some groups of children enter U.S. kindergarten classrooms already far less knowledgeable about the natural and social sciences than other groups of children. These preexisting general knowledge gaps in turn strongly predict general knowledge gaps in first grade, which in turn strongly predict science achievement gaps in third grade. Then, from third through eighth grades, experiencing lower reading and mathematics achievement is predictive of these science achievement gaps' persistence. Much of the sociodemographic science achievement gaps in the later elementary and middle school grades can be explained by modifiable

factors measured during the primary grades. Overall, the strongest contributors to science achievement gaps in the United States are general knowledge gaps that are already present at kindergarten entry. Therefore, interventions designed to address science achievement gaps in the United States may need to be implemented very early in children's development (e.g., by or around school entry if not earlier) so as to counteract the early onset of general knowledge gaps during the preschool and early elementary years.

However, these early-appearing gaps may be exacerbated by other modifiable factors. These include whether children also experience lower reading and mathematics achievement as they age and the racial-ethnic composition of the schools they attend, possibly due to attending lower-resourced schools. This suggests that policies and practices designed to address the nation's large and long-standing science achievement gaps may also need to be multifaceted and address lagging reading and mathematics achievement (Morgan et al., 2009), lower behavioral self-regulation (Duncan et al., 2007), and school racial segregation (Mickelson et al., 2013), which is currently increasing (Fiel, 2013). Consistent with other research (e.g., J.-S. Lee & Bowen, 2006), we find that higher-quality parenting is associated with greater academic achievement, although our results suggest that the unique association of parenting with science achievement is quite modest by the elementary and middle school grades.

Our findings also have theoretical implications. To date, whether the early achievement trajectories of children at risk are best described as stable, cumulative, or compensatory has been unclear. Our findings suggest that science achievement gaps begin to occur early in the school career and are largely stable as children age. However, our analyses provide some indication for both cumulative and compensatory models of achievement growth, although the extent to which these gaps increase or decrease over time is small. Consistent with prior work (e.g., Huang et al., 2014; Morgan et al., 2011), we observed that Black children often follow a cumulative trajectory in that they experience both initially lower and then somewhat slower science achievement growth. These gaps remain evident despite extensive statistical controls, although they are largely explained by the study's many other factors (e.g., reading and mathematics achievement, behavioral self-regulation, SES, and parenting). The science achievement gaps between lower- and higher-SES children are also consistent with a cumulative achievement growth model (McCoach et al., 2006). Consistent with a "threshold hypothesis," language-minority children initially display lower science achievement, but this is followed by greater achievement growth. However, their greater achievement growth was fully explained by the study's other modifiable factors. We failed to observe that school climate was uniquely associated with science achievement, possibly due to our use of a more extensive set of statistical controls (including children's time-varying reading and mathematics achievement) than those used in prior studies.

As has been shown often in reading as well as, increasingly, in mathematics (e.g., Cameron, Grimm, Steele, Castro-Schilo, & Grissmer, 2015; Morgan et al., 2011; Reardon, 2011), science achievement gaps begin early and then remain highly stable. Although some groups of children may display compensatory growth trajectories, their accelerated learning gains often are insufficient to close the initial gaps, particularly if they fail to receive continued and coordinated supports as they age (Protopapas, Sideridis, Mouzaki, & Simos, 2011). It may be that there is a common set of factors underlying achievement gaps that are domain general and so result in being more likely to struggle academically whether it be in reading, mathematics, science, or other content areas (Welsh, Nix, Blair, Bierman, & Nelson, 2010; Morgan, Farkas, Hillemeier, Hammer, & Maczuga, 2015).

Limitations

Our study has several limitations. First, the ECLS-K did not include a measure exclusively designed to evaluate children's science achievement during kindergarten and first grade. Instead, a measure of general knowledge was administered during these grades consisting of 50% science and 50% social studies items. Although children's general knowledge very strongly predicted their later science achievement, a measure consisting exclusively of science items in kindergarten and first grade would have allowed us to more directly evaluate the onset and over-time dynamics of science achievement gaps during this early time period. Second, and relatedly, the ECLS-K does not include attitudinal measures toward science. Children's attitudes, inclinations, and motivations toward science may have also predicted their achievement growth (Osborne, 2003) as well as related to children's feelings of identity (Archer et al., 2010). Finally, the ECLS-K data are nonexperimental. They do not allow for unambiguous causal inferences. Instead, the data allow only for hypothesis generation, while providing preliminary indication of factors uniquely associated with or predictive of science achievement gaps and so constituting potential targets of early intervention efforts.

Study's Contributions and Implications

Despite these limitations, our results help inform policies and interventions designed to address science achievement gaps in the United States. One set of findings informs economic and educational policymaking. The nation's long-standing science achievement gaps may be reduced by implementing policies to reduce economic inequality in the United States, including the intergenerational disadvantage of being raised in a low-SES family (Reardon, 2011). Prior studies have reported the importance of SES on children's achievement, but they pertain mostly to reading (Benson & Borman, 2010; Borman & Dowling, 2010; McCoach et al., 2006) or sometimes mathematics (A. Wang et al., 2013) but very rarely science achievement (Byrnes & Miller, 2007) and even less so the science achievement of elementary and middle school students (Sackes et al., 2011). Yet socioeconomic inequality has been increasing in the United States, including school socioeconomic segregation (Reardon & Bischoff, 2011). This is particularly likely to be experienced by children who are Black, Hispanic, and American Indian, who are about twice as likely to live in high-poverty areas as children who are White (Ann E. Casey Foundation, 2012). Our study suggests that if this demographic trend continues, one expected result may be even greater declines in the science achievement of U.S. schoolchildren. Our empirical findings are consistent with

theoretical accounts of science achievement gaps, including how modifiable child-, family- and school-level factors may help explain achievement gaps (Norman, Ault, Bentz, & Meskimen, 2001). Prior work has reported that greater racial and ethnic school segregation is predictive of lower reading or mathematics achievement (e.g., Benson & Borman, 2010; Berends & Penaloza, 2010; Mickelson et al., 2013). Our results support efforts to reduce racial and ethnic segregation in U.S. schools as a possible method for addressing achievement gaps as well as extend the field's knowledge base by establishing that the relation also pertains specifically to science achievement.

A second set of findings informs educational policymaking. Our results suggest that children's science achievement generally increases over time but that science achievement gaps initially evident by the elementary grades continue to occur as children age. Thus, policymakers and practitioners may need to increase the provision of early intervention efforts in science—particularly for at-risk populations—if science achievement gaps are to be narrowed or closed (Evans, 2005). That is, a potential way to reduce science achievement gaps may be by implementing policies providing children at risk with greater opportunities to learn about the natural and social sciences prior to or immediately following school entry, including by providing their parents with training in how to more effectively increase the children's readiness for schooling and/or through programs and policies that provide the children with greater access to informal learning opportunities, as often provided through high-quality child care and preschool. Substantial attention has already been focused on the need for such early childhood programs (e.g., Head Start, Early Reading First) to address emergent literacy and numeracy gaps during this early time period, as these gaps have been repeatedly established as highly predictive of later reading and mathematics achievement gaps (e.g., Chatterji, 2006; Downey, von Hippel, & Broh, 2004; Foster & Miller, 2007; Jordan, Kaplan, Ramineni, & Locuniak, 2009). Yet little attention has been given to the need to also address science achievement gaps during the preschool or early-elementary time periods (Tate et al., 2012). This is despite repeated and increasingly urgent calls for education policymakers and practitioners to work to prevent low science achievement from threatening the nation's economic competitiveness and the electorate's understanding of public policy issues (NASNAEIM, 2010, 2011). Time spent on science instruction is also currently declining in U.S. elementary schools, despite this instruction being associated with greater science achievement (Blank, 2013). Student interest in science begins to decline by age 11 (Osborne, 2003), suggesting again that interventions early in children's school careers may constitute "a key time for building interest" in science (Blank, 2013, p. 832). After the early-elementary time period, children, especially those who are at risk, may begin to internalize views of science as "hard" or for "eggheads" or mistakenly resulting from fixed ability (Archer et al., 2010), leading elementary and middle school teachers to confront attitudinal as well as academic barriers when trying to address science achievement gaps. Higher-quality elementary and middle school climates, which may help to increase reading and mathematics achievement (e.g., Bodovski et al., 2013; Bryk

et al., 2010), may be insufficient to address science achievement gaps. Instead, and because our findings indicate that science achievement gaps in the United States are largely stable and mostly explained by modifiable factors already occurring prior to or by school entry, multifaceted and coordinated early interventions providing increased opportunities for children who are at risk to learn about the natural and social sciences as well as about emergent literacy and numeracy, and to acquire attention and other self-regulatory behaviors, may need to be delivered prior to or immediately following school entry (Evans & Rosenbaum, 2008; Sylva, 2014). The provision of early and sustained interventions for children at risk may be necessary in order to close the nation's large science achievement gaps.

NOTE

REFERENCES

Allen, N. L., Carlson, J. E., & Zelenak, C. A. (1999). *The NAEP 1996 technical report* (NCES 1999-452). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Anderman, E. M. (1998). The middle school experience: Effects of the math and science achievement of adolescents with LD. *Journal of Learning Disabilities, 31*, 128–138. doi:10.1177/002221949803100203

Ann E. Casey Foundation. (2012). *Kids Count data snapshot on children living in American's high-poverty communities.* Retrieved from http://www.aecf.org/~/media/Pubs/Initiatives/KIDS%20COUNT/D/DataSnapshotonHighPovertyCommunities/KIDSCOUNTDataSnapshot_HighPovertyCommunities.pdf

Archer, L., Dewitt, J., Osborne, J., Dillion, J., Willis, B., & Wong, B. (2010). "Doing" science versus "being" a scientist: Examining 10/11-year-old schoolchildren's constructions of science through the lens of identity. *Science Education, 94*, 617–639. doi:10.1002/sce.20399

Ardasheva, Y., Tretter, T. R., & Kinny, M. (2012). English language learners and academic achievement: Revisiting the threshold hypothesis. *Language Learning, 62*, 769–812. doi:10.1111/j.1467-9922.2011.00652.x

Bacharach, V. R., Baumeister, A. A., & Furr, R. M. (2003). Racial and gender science achievement gaps in secondary education. *Journal of Genetic Psychology, 164*, 115–126. doi:10.1080/00221320309597507.

Bali, V. A., & Alvarez, R. M. (2004). The race gap in school achievement scores: Longitudinal evidence from a racially diverse school district. *Policy Studies Journal, 32*, 393–415. doi:10.1111/j.1541-0072.2004.00072.x

Batista, B., & Matthews, S. (2002). Integrated science and mathematics professional development programs. *School Science and Mathematics, 102*, 359–370.

Benson, J., & Borman, G. (2010). Family, neighborhood, and school settings across seasons: When do socioeconomic context and racial composition matter for the reading achievement growth of young children? *Teachers College Record, 112*, 1338–1390.

Berends, M., & Penaloza, R. (2010). Increasing racial isolation and test score gaps in mathematics: A 30-year perspective. *Teachers College Record, 112*, 978–1007.

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57*, 111–127. doi:10.1037/0003-066X.57.2.111

Blank, R. K. (2013). Science instructional time is declining in elementary schools: What are the implications for student achievement and closing the gap? *Science Education, 97*, 830–847. doi:10.1007/s11191-012-9537-6

Bodovski, K., Nahum-Shani, I., & Walsh, R. (2013). School climate and students' early mathematics learning: Another search for context effects. *American Journal of Education, 119*, 209–234.

Borman, G., & Dowling, M. (2010). Schools and inequality: A multilevel analysis of Coleman's equality of educational opportunity data. *Teachers College Record, 112*, 1201–1246.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53*, 371–399. doi:10.1146/annurev.psych.53.100901.135233

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago, IL: University of Chicago Press.

Byrnes, J. P., & Miller, D. C. (2007). The relative importance of predictors of math and science achievement: An opportunity–propensity analysis. *Contemporary Educational Psychology, 32*, 599–629. doi:10.1016/j.cedpsych.2006.09.002

Cameron, C. E., Grimm, K. J., Steele, J. S., Castro-Schilo, L., & Grissmer, D. W. (2015). Nonlinear Gompertz curve models of achievement gaps in mathematics and reading. *Journal of Educational Psychology, 107*, 789–804. doi:10.1037/edu0000009

Cascio, E. U., & Schanzenbach, D. W. (2013). *The impacts of expanding access to high-quality preschool education*. Brookings Papers on Economic Activity. Retrieved from http://www.brookings.edu/about/projects/bpea/papers/2013/fall-cascio-preschool-education

Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence from the Early Childhood Longitudinal Study (ECLS) kindergarten to first grade sample. *Journal of Educational Psychology, 98*, 489–507. doi:10.1037/0022-0663.98.3.489

Crosnoe, R. (2012). Studying the immigrant paradox in the Mexican-origin population. In C. G. Coll & A. K. Marks (Eds.), *The immigrant paradox in children and adolescents: Is becoming American a developmental risk?* (pp. 61–76). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/13094-003

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability ten years later. *Developmental Psychology, 33*, 934–945. doi:10.1037/0012-1649.33.6.934

Delen, E., & Bulut, O. (2011). The relationship between students' exposure to technology and their achievement in science and math. *Turkish Online Journal of Educational Technology, 10*, 311–317.

Downey, D. B., von Hippel, P. T., & Broh, B. A. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review, 69*, 613–635.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*, 1428–1446. doi:10.1037/0012- 1649.43.6.1428

Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon (Series Ed.), N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 1018–1095). Hoboken, NJ: Wiley.

Evans, G. W. (2005). Reframing the achievement gap. *Phi Delta Kappan, 86*, 582–589.

Evans, G. W., & Rosenbaum, J. (2008). Self-regulation and the income-achievement gap. *Early Childhood Research Quarterly, 23*, 504–514.

Feng, H., & Cartledge, G. (1996). Social skill assessment of inner city Asian, African, and European American students. *School Psychology Review, 25*, 228–239.

Fiel, J. E. (2013). Decomposing school resegregation: Social closure, racial imbalance, and racial isolation. *American Sociological Review, 78*, 828–848. http://dx.doi.org/10.1177/0003122413496252

Flores, A. (2007). Examining disparities in mathematics education: Achievement gap or opportunity gap? *High School Journal, 91*, 29–42.

Foster, W. A., & Miller, M. (2007). Development of the literacy achievement gap: A longitudinal study of kindergarten through third grade. *Language, Speech, and Hearing Services in Schools, 38*, 173–181. doi:10.1044/0161-1461(2007/018)

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*, 447–464.

Furlong, M. J., & Karno, M. (1995). Review of the Social Skills Rating System. In J. C. Conoley & J. C. Impara (Eds.), *Twelfth mental measurements yearbook* (pp. 697–969). Lincoln, NE: Buros Institute of Mental Measurement.

Gould, E., Mishel, L., & Shierholz, H. (2013). *Already more than a lost decade: Income and poverty trends continue to paint a bleak picture*. Washington, DC: Economic Policy Institute. Retrieved from http://www.epi.org/publication/lost-decade-income-poverty-trends-continue/

Gresham, F. M., & Elliot, S. N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.

Gustin, W. C., & Corazza, L. (1994). Mathematical and verbal reasoning as predictors of science achievement. *Roeper Review: A Journal on Gifted Education, 16*, 160–162. doi:10.1080/02783199409553564

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.

Hillemeier, M. M., Morgan, P. L., Farkas, G., & Maczuga, S. (2013). Quality disparities in child care for at-risk children: Comparing Head Start and non–Head Start settings. *Maternal and Child Health, 17*, 180–188.

Huang, F. L., Moon, T. R., & Boren, R. (2014). Are the reading rich getting richer? Testing for the presence of the Matthew effect. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 30*, 95–115. doi:10.1080/10573569.2013.789784

Jones, K. K., & Byrnes, J. P. (2006). Characteristics of students who benefit from high quality mathematics instruction. *Contemporary Educational Psychology, 31*, 328–343. doi:10.1016/j.cedpsych.2005.10.002

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45*, 850–867. doi:10.1037/a0014939

Kempert, S., Saalbach, H., & Hardy, I. (2011). Cognitive benefits and costs of bilingualism in elementary school students: The case of mathematical word problems. *Journal of Educational Psychology, 103*, 547–561. doi:10.1037/a0023619

Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in U.S. schools. *Educational Researcher, 35*, 3–12. doi:10.3102/0013189X035007003

Lee, J.-S., & Bowen, N. K. (2006). Parent involvement, cultural capital, and the achievement gap among elementary school children. *American Educational Research Journal, 43*, 193–218.

Lee, O., Deaktor, R. A., Hart, J. E., Cuevas, P., & Enders, C. (2005). An instructional intervention's impact on the science and literacy

achievement of culturally and linguistically diverse elementary students. *Journal of Research in Science Teaching, 42*, 857–887. doi:10.1002/tea.20071

Leppänen, U., Niemi, P., Aunola, K., & Nurmi, J. (2004). Development of reading skills among preschool and primary school pupils. *Reading Research Quarterly, 39*, 72–93. doi:10.1598/RRQ.39.1.5

Li, I., Onaga, E., Shen, P.-S., & Chiou, H.-H. (2009). Temperament characteristics and science achievement: A longitudinal study of elementary students in Taiwan. *International Journal of Science Achievement, 31*, 1175–1185.

Liu, X., & Whitford, M. (2011). Opportunities-to-learn at home: Profiles of students with and without reaching science proficiency. *Journal of Science Education Technology, 20*, 375–387. doi:10.1007/s10956-010-9259-y

Ma, X., & Wilkins, J. L. M. (2002). The development of science achievement in middle and high school: Individual differences and school effects. *Evaluation Review, 26*, 395–417. doi:10.1177/0193841X02026004003

Maerten-Rivera, J., Myers, N., Lee, O., & Penfield, R. (2010). Student and school predictors of high-stakes assessment in science. *Science Education, 94*, 937–962. doi:10.1002/sce.20408

Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). The persistence of preschool effects: Do subsequent classroom experiences matter? *Early Childhood Research Quarterly, 22*, 18–38. doi:10.1016/j.ecresq.2006.10.002

McClelland, M., Acock, A., & Morrison, F. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly, 21*, 471–490. doi:10.1016/j.ecresq.2006.09.003

McCoach, D. B., O'Connell, A. A., Reis, S. M., & Levitt, H. A. (2006). Growing readers: A hierarchical linear model of children's reading growth during the first 2 years of school. *Journal of Educational Psychology, 98*, 14–28. doi:10.1037/0022- 0663.98.1.14

McCoy, D. C., Roy, A. L., & Sirkman, G. M. (2013). Neighborhood crime and school climate as predictors of elementary school academic quality: A cross-lagged panel analysis. *American Journal of Community Psychology, 52*, 128–140.

Mickelson, R. A., Bottia, M. C., & Lambert, R. (2013). Effects of school racial composition on K–12 mathematics outcomes: A metaregression analysis. *Review of Educational Research, 83*, 121–158. doi:10.3102/0034654312475322

Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C., & Maczuga, S. (2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development, 86*, 1351–1370. doi:10.1111/cdev.12398

Morgan, P. L., Farkas, G., Tufis, P. S., & Sperling, R. S. (2008). Are reading and behavioral problems risk factors for each other? *Journal of Learning Disabilities, 41*, 417–436. doi:10.1177/0022219408321123

Morgan, P. L., Farkas, G., & Wu, Q. (2009) Five-year growth trajectories of kindergarten children with learning difficulties in mathematics. *Journal of Learning Disabilities, 42*, 306–321. doi:10.1177/0022219408331037

Morgan, P. L., Farkas, G., & Wu, Q. (2011). Kindergarten children's growth trajectories in reading and mathematics: Who falls increasingly behind? *Journal of Learning Disabilities, 44*, 472–488. doi:10.1177/0022219411414010

Morsy, L., & Rothstein, R. (2015). *Five social disadvantages that depress student performance: Why schools alone can't close achievement gaps.* Washington, DC: Economic Policy Institute. Retrieved from http://s2.epi.org/files/pdf/86987.pdf

Muller, P. A., Stage, F. K., & Kinzie, J. (2001). Science achievement growth trajectories: Understanding factors related to gender and racial-ethnic differences in precollege science achievement. *American Educational Research Journal, 38*, 981–1012. doi:10.3102/00028312038004981

Najarian, M., Pollack, J. M., & Sorongon, A. G. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K): Psychometric report for the eighth grade* (NCES 2009–002). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

National Academy of Sciences, National Academy of Engineering, & Institute of Medicine. (2010). *Rising above the gathering storm, revisited.* Washington, DC: National Academies Press.

National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. (2011) *Expanding underrepresented minority participation: America's science and technology talent at the crossroads.* Washington, DC: National Academies Press.

Norman, O., Ault, C. R., Bentz, B., & Meskimen, L. (2001). The Black–White "achievement gap" as a perennial challenge to urban science education: A sociocultural and historical overview with implications for research and practice. *Journal of Research in Science Teaching, 38*, 1101–1114. doi:10.1002/tea.10004

O'Reilly, T., & McNamara, D. S. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. *American Educational Research Journal, 44*, 161–196. doi:10.3102/0002831206298171

Osborne, J. (2003). Attitudes toward science: A review of the literature and its implications. *International Journal of Science Education, 25*, 1049–1079. doi:10.1080/0950069032000032199

Pollack, J. M., Atkins-Burnett, S., Najarian, M., & Rock, D.A. (2005). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS–K): Psychometric report for the fifth grade* (NCES 2006–036).Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Pong, S. L., Dronkers, J., & Hampden-Thompson, G. (2003). Family policies and children's school achievement in single- versus two-parent families. *Journal of Marriage and Family, 65*, 681–699. doi:10.1111/j.1741-3737.2003.00681.x

Protopapas, A., Sideridis, G. D., Mouzaki, A., & Simos, P. G. (2011). Matthew effects in reading comprehension: Myth or reality? *Journal of Learning Disabilities, 44*, 402–420. doi:10.1177/0022219411417568

Raghunathan, T. E., Lepkowski, J. M., Van Hoewyk, J., & Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology, 27*, 85–95.

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79*, 375–394. doi:10.1111/j.1467- 8624.2007.01131.x

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* Thousand Oaks, CA: Sage.

Reardon, S. F. (2011). The widening academic achievement gap between rich and poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 91–115). New York, NY: Russell Sage Foundation.

Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation. *American Journal of Sociology, 116*, 1092–1153. doi:10.1086/657114

Reardon, S. F., & Galindo, C. (2009). The Hispanic–White achievement gap in math and reading in the elementary grades. *American Educational Research Journal, 46*, 853–891. http://dx.doi.org/10.3102/0002831209333184

Reynolds, A. J., & Walberg, H. J. (1992). A structural model of science achievement and attitude: An extension to high school. *Journal of Educational Psychology*, *84*, 371–382. doi:10.1037/0022-0663.84.3.371

Ricciardelli, L. A. (1992). Bilingualism and cognitive-development in relation to threshold theory. *Journal of Psycholinguistic Research*, *21*, 301–316. doi:10.1007/BF01067515

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*. Working Paper Series. Jessup, MD: ED Pubs.

Sackes, M., Trundle, K. C., Bell, R. L., & O'Connell, A. A. (2011). The influence of early science experience in kindergarten on children's immediate and later science achievement: Evidence from the Early Childhood Longitudinal Study. *Journal of Research in Science Teaching*, *48*, 217–235. doi:10.1002/tea.20395

Scarborough, H. S. (1998). Predicting the future achievement of second graders with reading disabilities: Contributions of phonemic awareness, verbal memory, rapid naming, and IQ. *Annals of Dyslexia*, *68*, 115–136. doi:10.1007/s11881-998- 0006-5

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford, UK: Oxford University Press.

Skibbe, L. E., Grimm, K. J., Stanton-Chapman, T., Justice, L. M., Pence, K. L., & Bowles, R. P. (2008). Reading trajectories of children with language difficulties from preschool through fifth grade. *Language, Speech & Hearing Services in Schools*, *39*, 475–486. doi:10.1044/0161-1461(2008/07-0016)

Slykhuis, D., & Park, J. (2006). Correlates of achievement with online and classroom-based MBL physics activities. *Journal of Computers in Mathematics and Science Teaching*, *25*, 147–163.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–407.

Stewart, E. B. (2008). School structural characteristics, student effort, peer associations, and parental involvement: The influence of school- and individual-level factors on academic achievement. *Education and Urban Society*, *40*, 179–204.

Strand, S. (2014). School effects and ethnic, gender, and socio-economic gaps in educational achievement at age 11. *Oxford Review of Education*, *40*, 223–245. doi:10.1080/03054985.2014.891980

Sylva, K. (2014). The role of families and pre-school in educational disadvantage. *Oxford Review of Education*, *40*, 680–695.

Tach, L., & Farkas, G. (2006). Learning-related behaviors, cognitive skills, and ability grouping when schooling begins. *Social Science Research*, *35*, 1048–1079. doi:10.1016/j.ssresearch.2005.08.001

Tate, W. F., Jones, B. D., Thorne-Wallington, E., & Hogrebe, M. C. (2012). Science and the city: Thinking geospatially about opportunity to learn. *Urban Education*, *47*, 399–433. doi:10.1177/0042085911429974

U.S. Census Bureau. (2012a). *Children below poverty level by race and Hispanic origin*. Retrieved from http://www.census.gov/compendia/statab/2012/tables/12s0712.pdf

U.S. Census Bureau. (2012b). *Most children younger than age 1 are minorities*. Retrieved from http://www.census.gov/newsroom/releases/archives/population/cb12-90.html

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2015). *National Assessment of Education Progress, science 2011, national results*. Retrieved from http://www.nationsreportcard.gov/science_2011/

U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics. (2000). *NAEP 1999 trends in academic progress: Three decades of student performance* (NCES 2000–469). Retrieved from http://nces.ed.gov/nationsreportcard/pdf/main1999/2000469.pdf

Von Secker, C. (2004). Science achievement in social contexts: Analysis from the National Assessment of Educational Progress. *Journal of Educational Research*, *98*, 67–78. doi:10.3200/JOER.98.2.67-78

Von Secker, C., & Lissitz, R. W. (1999). Estimating the impact of instructional practices on student achievement in science. *Journal of Research in Science Teaching*, *36*, 1110–1126. doi:10.1002/(SICI)1098-2736(199912)36:10%3C1110::AID-TEA4%3E3.0.CO;2-T

Walberg, H. J., & Tsai, S. (1983). Matthew effects in education. *American Educational Research Journal*, *20*, 359–373. http://dx.doi.org/10.3102/00028312020003359

Wang, A. H, Shen, F., & Byrnes, J. P. (2013). Does the opportunity-propensity framework predict the early mathematics skills of low-income pre-kindergarten children? *Contemporary Educational Psychology*, *38*, 259–270. doi:10.1016/j.cedpsych.2013.04.004

Wang, J., & Staver, J. R. (2001). Examining relationships between factors of science education and student career aspiration. *Journal of Educational Research*, *94*, 312–319. doi:10.1080/00220670109598767

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology*, *102*, 43–53.

Woodcock, R. W., McGrew, K. S., & Werder, J. K. (1994). *Woodcock-McGrew-Werder Mini-Battery of Achievement*. Itasca, IL: Riverside.

Young, D. J., Reynolds, A. J., & Walberg, H. J. (1996). Science achievement and educational productivity. *Journal of Educational Research*, *89*, 272–278. doi:10.1080/00220671.1996.9941328

AUTHORS

PAUL L. MORGAN, PhD, is an associate professor of education, Department of Education Policy Studies, Pennsylvania State University, 310E Rackley Building, University Park, PA 16802; *paulmorgan@psu.edu*. His research examines factors that may contribute to academic or behavioral struggles in school.

GEORGE FARKAS, PhD, is a professor of education at the University of California–Irvine, School of Education, 3200 Education, Irvine, CA 92697; *gfarkas@uci.edu*. His research focuses on educational inequality and how it can be reduced.

MARIANNE M. HILLEMEIER, PhD, MPH, is a professor of health policy and administration and demography at the Pennsylvania State University, 604E Ford Building, University Park, PA 16802; *mmh18@psu.edu*. Her research focuses on socioeconomic, racial/ethnic, and geographic disparities in children's health and well-being.

STEVE MACZUGA, MS, is a research programmer at the Population Research Institute, Pennsylvania State University, 601 Oswald Tower, University Park, PA 16802; *sam21@psu.edu*. His research focuses on regression analysis of longitudinal surveys of childhood education and health.