Journal of Applied Psychology

Explaining the Black–White Gap in Cognitive Test Scores: Toward a Theory of Adverse Impact

Jonathan M. Cottrell, Daniel A. Newman, and Glenn I. Roisman

Online First Publication, April 13, 2015. http://dx.doi.org/10.1037/apl0000020

Journal of Applied Psychology

0021-9010/15/$12.00 http://dx.doi.org/10.1037/apl0000020

Explaining the Black–White Gap in Cognitive Test Scores: Toward a Theory of Adverse Impact

Jonathan M. Cottrell and Daniel A. Newman
University of Illinois at Urbana-Champaign

Glenn I. Roisman
University of Minnesota

In understanding the causes of adverse impact, a key parameter is the Black–White difference in cognitive test scores. To advance theory on why Black–White cognitive ability/knowledge test score gaps exist, and on how these gaps develop over time, the current article proposes an inductive explanatory model derived from past empirical findings. According to this theoretical model, Black–White group mean differences in cognitive test scores arise from the following racially disparate conditions: family income, maternal education, maternal verbal ability/knowledge, learning materials in the home, parenting factors (maternal sensitivity, maternal warmth and acceptance, and safe physical environment), child birth order, and child birth weight. Results from a 5-wave longitudinal growth model estimated on children in the NICHD Study of Early Child Care and Youth Development from ages 4 through 15 years show significant Black–White cognitive test score gaps throughout early development that did not grow significantly over time (i.e., significant intercept differences, but not slope differences). Importantly, the racially disparate conditions listed above can account for the relation between race and cognitive test scores. We propose a parsimonious 3-Step Model that explains how cognitive test score gaps arise, in which race relates to maternal disadvantage, which in turn relates to parenting factors, which in turn relate to cognitive test scores. This model and results offer to fill a need for theory on the etiology of the Black–White ethnic group gap in cognitive test scores, and attempt to address a missing link in the theory of adverse impact.

Keywords: adverse impact, cognitive ability/knowledge, race, personality, achievement gap

In the study of adverse impact, racial group differences in cognitive ability/knowledge test scores represent a classic problem (see Equal Employment Opportunity Commission, 1978; Goldstein, Scherbaum, & Yusko, 2010; *Griggs v. Duke Power Co.*, 1971; Outtz, 2010; Schmitt & Quinn, 2010; Zedeck, 2010). More specifically, personnel selection practices are a key mechanism by which individuals from different racial backgrounds gain access to jobs. Nonetheless, a major empirical tension has plagued the process of hiring and admissions as pertains to adverse impact/diversity and job performance (Outtz, 2010; Sackett, Schmitt, Ellingson, & Kabin, 2001). On the one hand, cognitive tests robustly predict job performance across many different job types, and are considered the predictor of choice for achieving maximal job performance (Hunter & Hunter, 1984; Schmidt & Hunter, 1998, 2004). On the other hand, these cognitive tests also show large Black–White subgroup differences, with an average Cohen's *d*, or standardized mean difference, of 1.0 (Roth, Bevier, Bobko, Switzer, & Tyler, 2001; Sackett & Shen, 2010; as a point of reference, the Black–White gap in cognitive tests is nearly three times as large as the Black–White gap in job performance itself—see McKay & McDaniel, 2006). In other words, using cognitive tests for hiring purposes will tend to exclude African American job applicants (Bobko, Roth, & Potosky, 1999; Newman, Jacobs, & Bartram, 2007; Sackett & Ellingson, 1997; Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997), and can lead to large-scale race disparities in occupational attainment.

Many possibilities have been suggested for how to resolve this tension between the large criterion validity of, and large ethnic subgroup differences on, cognitive tests (for a review, see Ployhart & Holtz, 2008). In the short term, selection systems can give greater weight to noncognitive tests, such as personality inventories, interviews, assessment centers, or situational judgment tests. Although these noncognitive instruments do in fact show less adverse impact potential (Bobko & Roth, 2013), they also are less valid predictors of job performance, and would therefore have lower monetary utility for firms if used as a substitute for cognitive tests (Ployhart & Holtz, 2008; Schmidt & Hunter, 1998). Optimal tradeoffs between validity and diversity in hiring continue to be explored (e.g., De Corte, Lievens, & Sackett, 2007; De Corte, Sackett, & Lievens, 2010; Wee, Newman, & Joseph, 2014).

Beneath these practical concerns lies a more fundamental problem: Adverse-impact researchers lack a comprehensive understanding of the *origins* of the race gap in cognitive test scores. That is, we lack a parsimonious theory of adverse impact (Outtz &

Jonathan M. Cottrell, Department of Psychology, University of Illinois at Urbana-Champaign; Daniel A. Newman, Department of Psychology and School of Labor & Employment Relations, University of Illinois at Urbana-Champaign; Glenn I. Roisman, Institute of Child Development, University of Minnesota.

Correspondence concerning this article should be addressed to Jonathan M. Cottrell, Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, IL 61820. E-mail: cottrell601@gmail.com

Newman, 2010). To understand how adverse impact develops, we must understand how cognitive ability/knowledge develops by examining childhood and familial factors that play a role in its development.

Although adverse impact is a long-established, core topic in the field of Industrial/Organizational (I/O) psychology, some readers might reasonably question how a study of childhood and familial factors that lead to Black–White gaps in cognitive test scores contributes meaningfully to organizational theory. In response, we note that the current study can help to develop a theory of adverse impact. To elaborate, the phenomenon of adverse impact is defined by a pattern among three relations (Outtz & Newman, 2010): (a) the relation of cognitive tests to job performance (Hunter & Hunter, 1984; Schmidt & Hunter, 1998), (b) the relation of job performance to race (McKay & McDaniel, 2006; Roth, Huffcutt, & Bobko, 2003), and (c) the relation of race to cognitive test scores (Roth et al., 2001). Of these three defining relations in adverse impact, two have been theoretically explained, whereas the third—conspicuously—has not. The field of I/O psychology has a long-standing and fairly adequate theoretical explanation for the association between cognitive tests and job performance (i.e., cognitive tests relate to job performance due to the acquisition of job knowledge; see Schmidt & Hunter, 2004; Schmidt, Hunter, & Outerbridge, 1986), and I/O psychologists also have a longstanding and fairly adequate theoretical explanation for the relationship between race and job performance (i.e., race differences in job performance can be almost fully explained by race differences in cognitive test scores and job experience; see Schmidt & Hunter, 1981; Waldman & Avolio, 1991). What is notably missing from our field's current theory of adverse impact is an adequate explanation for the sizable race differences in cognitive test scores.

Indeed, in Roth et al.'s (2001) classic article on race differences in cognitive test scores, they recommend that the two primary "avenues for future research" are (a) "testing strategies," and (b) "reasons for ethnic group differences" (p. 324). The latter research avenue echoes a sentiment expressed decades earlier by Schmidt and Hunter (1981), who said,

> The average ability and cognitive skill differences between groups are directly reflected in job performance, and thus are real. They are *not* created by the tests. *We do not know what all the causes of these differences are* [emphasis added], how long they will persist, or how best to eliminate them. (p. 1131)

As stated by the Society for I/O Psychology (SIOP) Adverse Impact Reduction Research Initiative and Action group, "Adverse impact (AI) has been a recurring problem for decades, and although some findings have not only helped us to understand the problem and create modest solutions, we are still without a viable theory" (SIOP, 2013). The current article will attempt to contribute information about the third and missing piece of the theory of adverse impact, by offering one possible explanation for the relation between race and cognitive test scores.

Overall, the current research will try to make four contributions toward improving our understanding of how adverse impact develops. First, we track the Black–White gap in cognitive test scores longitudinally across five time points from ages 4 to 15 years, to show that there exists an intercept gap but no race difference in slopes/growth curves. This result suggests that the gap already exists at 4 years of age, and remains somewhat stable in magnitude thereafter. Second, we offer an interdisciplinary review of studies from economics, educational psychology, and developmental psychology that have tried to explain this gap empirically, and use this review to construct a list of the mechanisms through which the gap originates. Third, we demonstrate that this set of explanatory covariates (e.g., parenting factors, socioeconomic status [SES], birth order, learning materials in the home) can together explain the Black–White cognitive test score gap. Fourth and finally, we propose a *3-Step Model* of race gaps in cognitive scores, to provide a more parsimonious explanation for why the Black–White gap arises.

Previous Attempts to Explain the Black–White Cognitive Test Score Gap

The Black–White gap in test scores is widely known, but not widely understood. Despite the fact that the Black–White gap in cognitive test scores has been discussed for over 90 years (e.g., Popenoe, 1922), surprisingly few studies have attempted to *empirically* test integrated theoretical models that explain this gap. Nonetheless, there have been a few attempts to quantitatively explain the Black–White test score gap (as summarized in Table A1 of Appendix A). Note that Table A1 does not include articles that only quantify the size of the gap (e.g., Roth et al., 2001), nor does Table A1 include articles that provide only theoretical explanations for the Black–White gap, with no data (e.g., Brooks-Gunn & Markman, 2005; Garcia Coll et al., 1996).

Explanatory Variables Identified in Past Research on the Cognitive Test Gap

Table 1 enumerates the covariates that have repeatedly been found in past research to uniquely explain part of the Black–White cognitive test score gap. In other words, these are the explanatory concepts whose unique statistical significance has been replicated (i.e., been found in more than one past study). These concepts, each of which describes a set of disparate conditions between Black children and White children, are birth order, maternal verbal ability/knowledge, learning materials in the home, parenting factors (maternal sensitivity, warmth and acceptance, physical environment), birth weight, and SES. Each of these explanatory concepts is hypothesized to correlate both with cognitive test scores and with race. After describing these explanatory concepts, we propose a 3-Step Model, which offers a more parsimonious view of how the explanatory variables relate to each other and to the test gap.

Birth order. Earlier born children have higher cognitive test scores than their later born siblings (Black, Devereux, & Salvanes, 2005; Booth & Kee, 2009), likely because firstborn and early born children receive (a) more parental time and attention (Behrman & Taubman, 1986), (b) a larger share of parental income spent on the child's development (Becker & Lewis, 1973), (c) a greater share of educational resources (e.g., books) divided among fewer siblings (Booth & Kee, 2009), (d) an intellectual environment that is less "diluted" by the simple language of other young children (Zajonc & Bargh, 1980), and (e) opportunities to serve as tutors and surrogate caregivers for younger children, especially in large families (Bargh & Schul, 1980; Heiland, 2009; Zajonc & Markus, 1975; Zajonc & Mullally, 1997; Zajonc & Sulloway, 2007).

Table 1
Replicated Covariates That Partly Explained the Black-White Gap in Cognitive Test Scores (Statistically Significant Across Multiple Samples)

Covariates	Number of samples where supported	References
Birth order/firstborn child	3	Yeung & Pfeiffer (2009) Cohorts 1 and 2; Burchinal et al. (2011)
Mother's verbal ability/knowledge	4	Brooks-Gunn et al. (2003, IHDP); Mandara et al. (2009); Yeung & Pfeiffer (2009) Cohorts 1 and 2
Learning materials	3	Brooks-Gunn et al. (2003, IHDP and NLSY-CS); Fryer & Levitt (2004/2006)
Maternal sensitivity/home warmth, maternal acceptance, physical environment	3	Brooks-Gunn et al. (2003, IHDP and NLSY-CS); Burchinal et al. (2011)
Birth weight	4	Brooks-Gunn et al. (2003, NLSY-CS); Fryer & Levitt (2004, 2006); Yeung & Pfeiffer (2009) Cohorts 1 and 3
SES	6	Brooks-Gunn et al. (2003, IHDP and NLSY-CS); Fryer & Levitt (2004/2006); Mandara et al. (2009); Yeung & Pfeiffer (2009) Cohorts 1 and 2; Burchinal et al. (2011)
• SES composite	4	Fryer & Levitt (2004, 2006); Mandara et al. (2009); Yeung & Pfeiffer (2009) Cohorts 1 and 2
• Income	3	Brooks-Gunn et al. (2003, IHDP and NLSY-CS); Burchinal et al. (2011)
• Maternal education	3	Brooks-Gunn et al. (2003, IHDP and NLSY-CS); Burchinal et al. (2011)

Note. SES = socioeconomic status; IHDP = Infant Health and Development Program; NLSY-CS = National Longitudinal Study of Youth - Child Supplement.

Black families in the United States, on average, have 25% more (of their own) children under 18 years of age in their households (M = 1.01 children; N = 9.3 million Black households) compared with White families (M = 0.81 children; N = 64 million White households; United States Census Bureau, 2010). As such, Black children tend, on average, to have a greater number of older siblings.

Maternal verbal ability/knowledge. The verbal ability/knowledge of one's mother has been associated with the cognitive ability/knowledge of children (Bennett, Bendersky, & Lewis, 2008). Children of high-ability, well-educated mothers tend to learn longer, more complex, and a larger number of words at a young age, likely due to a greater variety and complexity of words used by their mothers (Dollaghan et al., 1999). Schady (2011) showed that mothers with higher vocabulary levels, as measured by the Peabody Picture Vocabulary Test (Dunn & Dunn, 1981), had children with more advanced vocabulary and higher scores on tests of memory and visual integration. Relatedly, mothers with higher verbal test scores generally possess more self-esteem, academic aptitude, and higher expectations for both themselves and their children (Magnuson, 2007).

Maternal verbal ability/knowledge is also related to race. We assert that the robust Black–White cognitive test score gap (d = 1.0; Roth et al., 2001) generalizes to racial differences in verbal ability among mothers. Consistent with this generalization, previous studies have found a negative association between race and maternal verbal ability/knowledge test scores (e.g., d = 1.11, Mandara, Varner, Greene, & Richman, 2009; d = 1.35, Yeung & Pfeiffer, 2009).

Learning materials. Learning materials (e.g., books, puzzles) make the home environment educationally stimulating (Watson, Kirby, Kelleher, & Bradley, 1996), and are positively related to cognitive test scores (Linver, Brooks-Gunn, & Kohen, 2002). The learning materials subscale of the Home Observation for the Measurement of the Environment (HOME; Caldwell & Bradley, 1984) has been related to vocabulary, math, and reading tests, especially for younger children (Bradley, Corwyn, Burchinal, McAdoo, & Garcia Coll, 2001b).

There also exist large Black–White differences in learning materials in the home (d = 1.17, Bradley & Caldwell, 1984; d = 1.23, Brooks-Gunn, Klebanov, Smith, Duncan, & Lee, 2003 Infant Health and Development Program [IHDP]; d = 1.05, Thompson et al., 1998). One possible reason for this is that Black families tend to be poorer than White ones (as we describe later), making it harder to acquire learning materials (Linver et al., 2002). It is also possible that a parent's perception of the norms for how many books, puzzles, and other learning materials a young child needs stems partly from the parent's own experience as a child, regardless of current income (e.g., learning materials *uniquely* explain part of the Black–White gap; Brooks-Gunn et al., 2003; Fryer & Levitt, 2004, 2006).

Parenting factors: Maternal sensitivity, maternal warmth and acceptance, and physical environment. We differentiate learning materials (described above) from aspects of a child's environment related to mother's sensitivity and providing a secure and welcoming environment. We believe these parenting climate factors (maternal sensitivity, warmth and acceptance, and physical environment) are all related to the provision of a safe and responsive home for children, which fosters cognitive development.

Maternal sensitivity is generally defined as "a mother's ability to perceive and interpret accurately her infant's signals and communications and then respond appropriately" (Ainsworth, Bell, & Stayton, 1974, p. 127). This construct tends to be measured in the context of children doing difficult tasks, in which mothers' behavior in supporting and scaffolding children's task accomplishment is coded. In essence, maternal sensitivity often captures mothers' use of cognitive stimulation in a positive emotional climate (De Wolff & van IJzendoorn, 1997; Fraley, Roisman, & Haltigan, 2013). Early maternal sensitivity predicts academic and cognitive achievement (Estrada, Arsenio, Hess, & Holloway, 1987; Lemelin, Tarabulsy, & Provost, 2006; Page, Wilhelm, Gamble, & Card, 2010; Raby, Roisman, Fraley, & Simpson, in press; Stams, Juffer,

& van IJzendoorn, 2002). Using a similar rationale, the extent to which parents are accepting, act warmly toward their kids, and do not punish them harshly for mistakes has been related to cognitive test scores (Bradley et al., 1989; Brooks-Gunn et al., 2003). This effect of maternal acceptance and warmth has also been shown for math and reading scores, particularly among Black participants and poor White participants (Bradley, Corwyn, Burchinal, McAdoo, & Garcia Coll, 2001).

Finally, living in a home that is not overcrowded, is safe, and is relatively bright (i.e., physical environment) is positively related to academic achievement test scores (Bradley, Caldwell, Rock, Hamrick, & Harris, 1988). Physical environment might well be included in a general syndrome of parenting factors—along with maternal sensitivity and warmth—that share the common theme of facilitating and supporting children's cognitive development in the face of challenging tasks and learning scenarios. Consistent with this view, evidence shows that physical environment is strongly related to scores on maternal acceptance scales (r = .52) from home observation measures (Bradley et al., 1992). This suggests that parents who provide safe and healthy physical environments for children are also generally warm and accepting of their children, and are also involved in helping their children develop cognitively (Bradley et al., 1992).

Previous research has reported sizable Black–White gaps in maternal sensitivity (e.g., d = .94, Dotterer, Iruka, & Pungello, 2012; d = .44, Huang, Lewin, Mitchell, & Zhang, 2012; d = .63, Pungello, Iruka, Dotterer, Mills-Koonce, & Reznick 2009), maternal warmth and acceptance (e.g., d = .49, Bradley & Caldwell, 1984; d = .77, Brooks-Gunn et al., 2003 National Longitudinal Study of Youth - Child Suppplement), and physical environment (e.g., d = .68, Bradley & Caldwell, 1984; d = .41, Thompson et al., 1998). One possible explanation for this phenomenon is that discrimination against Black individuals may contribute to Black mothers' anxiety and depression, which could reduce the quality of the mother–child relationship (Pungello et al., 2009). Additionally, some scholars have posited that Black parents' having to cope with prejudice, as well as the tendency to live in more impoverished neighborhoods, might contribute to Black–White differences in parenting practices and home conditions (Bradley & Caldwell, 1984; Bradley, Corwyn, McAdoo, & Garcia Coll, 2001).

Birth weight. A recent meta-analysis showed that children of very low birth weight had significantly reduced volumes of the total brain, gray matter, white matter, cerebellum, hippocampus, and corpus callosum—all of which are related to lower cognitive test scores (de Kieviet, Zoetebier, van Elburg, Vermeulen, & Oosterlaan, 2012; Dezoete, MacArthur, & Tuck, 2003; see also Torche & Echevarría, 2011). Even low-birth-weight children without major neurosensory disorders such as cerebral palsy still have significantly lower cognitive test scores than children of normal birth weight (Taylor, Klein, Minich, & Hack, 2000).

Previous studies have shown that African American children have significantly lower birth weight than White children (e.g., d = .48, Lhila & Long, 2012; d = .33, Yeung & Conley, 2008). Some proposed reasons are that White mothers are often in higher socioeconomic conditions and are physically healthier than Black mothers (Lhila & Long, 2012).

SES: Income and maternal education. SES has been found to correlate with cognitive test scores, as well as with college grade point average (Sackett, Kuncel, Arneson, Cooper, & Waters, 2009). Family income, often used as an indicator of SES, can predict IQ as early as age 2 years (Klebanov, Brooks-Gunn, McCarton, & McCormick, 1998), and at later ages (Brooks-Gunn et al., 2003; Fryer & Levitt, 2004). Maternal education, another frequently used indicator of SES, also predicts cognitive test scores (Dollaghan et al., 1999; Magnuson, 2007; Schady, 2011). Further, being a low-income student in a high-income school negatively predicts science and math achievement scores (Crosnoe, 2009).

Poor families have fewer children's books and are likely to live in more dangerous neighborhoods, which limits access to educational resources (Duncan & Magnuson, 2005) and impairs learning and language stimulation (Watson et al., 1996; Yeung, Linver, & Brooks-Gunn, 2002). Additionally, more educated mothers have higher educational aspirations for their children, and the children achieve greater gains in reading and math skills (Magnuson, 2007), while acquiring more spontaneous language and words (Dollaghan et al., 1999). Low-SES parents have less time to spend with children, and experience stress from lack of resources, which can influence parents' warmth and sensitivity (Greenman, Bodovski, & Reed, 2011; McLoyd, 1990; Mistry, Benner, Biesanz, Clark, & Howes, 2010). Children of low-SES families are also more likely to experience growth retardation, learning disability, and child abuse (Brooks-Gunn & Duncan, 1997), which harm cognitive development.

Income and maternal education are related to race. Fifty-one percent of Black families earn less than $40,000 annually, whereas only 30% of White families have incomes less than $40,000 (United States Census Bureau, 2009a). The median income of Black individuals ($38,409) is $24,000 less than that of White individuals ($62,545; United States Census Bureau, 2009b). Additionally, as of 2010, 29.9% of White females had obtained at least a 4-year college degree, in contrast with only 21.4% of African American females (i.e., White females were 40% more likely to have a 4-year degree; United States Census Bureau, 2012a). African Americans are also much more likely to be enrolled in poor and overcrowded schools (Condron, 2009), implying that race differences in the amount of maternal education might actually understate differences in the quality of education.

3-Step Model of the Race Gap

The current article offers a novel theoretical explanation for race gaps in cognitive test scores, by presenting the first attempt to integrate all explanatory variables identified by past empirical research (see Table 1). Beyond simply examining the explanatory power of the combination of explanatory variables described above, it is possible to offer a more parsimonious model, by specifying a particular pattern of relations among these explanatory variables. To this end, we now introduce a 3-Step Model of the race gap in cognitive test scores (see Figure 1).

According to the 3-Step Model (Figure 1, Model A), race is related to cognitive test scores due to a sequential process in which: (Step 1) race gives rise to group differences in a set of concepts known as *maternal advantage* (i.e., income, maternal education, and maternal verbal ability/knowledge), (Step 2) maternal advantage leads to *parenting factors*, including maternal sensitivity, acceptance, physical environment, learning materials in

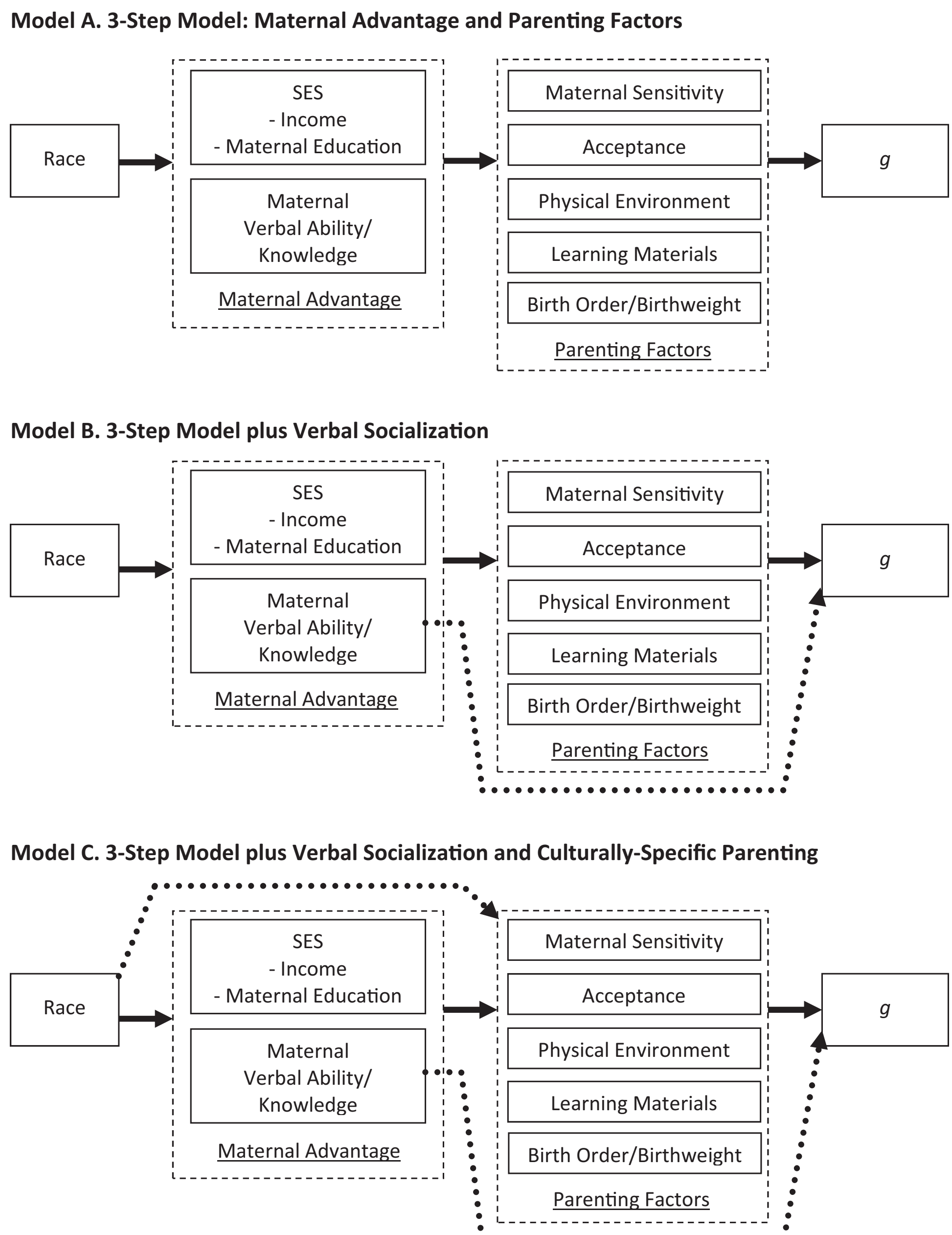

Figure 1. Theoretical models of the race/cognitive-test-score relationship.

the home, birth weight, and birth order; and finally, (Step 3) parenting factors in turn promote cognitive ability/knowledge.

Step 1: Race Relates to Maternal Advantage

To theoretically explain why race contributes to maternal advantage, we must begin by attempting to define what *race* is. According to Winant (2000),

> Race can be defined as a concept that symbolizes and signifies sociopolitical conflicts and interests in reference to different types of human bodies. Although the concept of race appeals to biologically based human characteristics (phenotypes), selection of these particular human features for purposes of racial signification is always and necessarily a social and historical process. (p. 172).

We also direct the reader to Omi and Winant (1994), for a description of how racial categories subjectively emerge. Thus, race, even when measured via simple self-categorization, is a complex construct that reflects aspects of cultural experience and identity (Phinney, 1996), as well as sociopolitical history (Feagin, 2006).

In the United States, the concept of race implies a history of housing segregation, educational segregation (which is maintained due to housing segregation in regions in which school access is based upon where one lives), and occupational segregation. In particular, this segregation is a result of three major factors spanning several centuries in the United States (Feagin, 2006). First, African Americans were legally enslaved from the early 1600s until 1865. Second, educational and employment segregation based on race were legal until the Civil Rights Act of 1964. Finally, post-World War II legislation excluded farm workers and maids, most of whom were Black, from minimum-wage protections, work-hour regulations, and unions (Katznelson, 2005). The combination of these and other factors resulted in major educational, occupational, and income gaps between Whites and Blacks, which can, in large part, explain the contemporary problem of adverse impact (Newman, Hanges, & Outtz, 2007). Race-based occupational segregation begets income differences, race-based educational segregation begets racial differences in maternal education and maternal verbal ability/knowledge, and race-based housing segregation begets educational and occupational segregation. Many books and articles have been written on these topics (e.g., Cutler, Glaeser, & Vigdor, 1999; Feagin, 2006; Higginbotham, 1980, 1996; Myrdal, 1944), and we simply summarize this vast literature by hypothesizing Step 1 in the 3-Step Model (see Figure 1).

Step 2: Maternal Advantage Relates to Parenting Factors

Higher income gives parents access to safer physical environments and the opportunity to buy more learning materials (Duyme, Dumaret, & Tomkiewicz, 1999; Watson et al., 1996). Additionally, Fox, Platz, and Bentley (1995) found that mothers with higher income and higher education tended to be more nurturing and to use less discipline, possibly due to growing up in a more nurturing environment themselves. Maternal education also gives parents more motivation to provide learning materials (Klebanov, Brooks-Gunn, & Duncan, 1994; Linver et al., 2002), more knowledge of the dangers of unsafe physical environments (Klebanov et al., 1994), and access to better jobs (Linver et al., 2002), which allow for safer and more stimulating home environments and better access to health care (both prenatal and birth control). Similarly, more verbally skilled and more educated parents can create more stimulating and protective home experiences (Bacharach & Baumeister, 1998; Brooks-Gunn & Markman, 2005). Finally, greater maternal education leads to smaller family size/lower birth order (Travis & Kohli, 1995) and higher birth weights (Rauh, Achenbach, Nurcombe, Howell, & Teti, 1988). In all, these explanations give rise to Step 2 in our hypothesized theoretical model—maternal advantage leads to more sensitive parenting, safer physical environments, more acceptance, provision of more learning materials, higher birth weights, and earlier birth order.

Step 3: Parenting Factors Relate to Cognitive Scores

Finally, for explaining the relation between parenting factors and *g*, we have already discussed these mechanisms previously, in the sections in which we introduced each parenting factor. Altogether, Steps 1, 2, and 3 combine to yield the 3-Step Model of race gaps in cognitive scores. This model offers an integrative theoretical account of why race gaps in cognitive achievement appear in U.S. society.

Alternative Model Specifications

Up to this point, we have proposed the 3-Step Model (Figure 1, Model A), which is a three-step *full mediation* model (James, Mulaik, & Brett, 2006). That is, no direct effects are specified between race and parenting variables, or between race and *g*; in addition, no direct effects are specified between maternal advantage variables and *g*. The rationale for the full mediation aspect of the 3-Step Model (Figure 1, Model A) consists of three simple assertions: (a) social and historical factors of de facto racial segregation in education and occupations have directly led to race gaps in maternal advantage (Feagin, 2006), (b) parenting factors are the major mechanism through which maternal advantage comes to affect child cognitive development (see our rationale for Step 2 and Step 3 above), and (c) racial differences in parenting factors are fully explained by maternal advantage, because lack of resources, education, and verbal ability limit parents' capacity, skill, and knowledge to provide a parenting environment that is sensitive, accepting, learning-oriented, and safe.

Nonetheless, there still exist the alternative possibilities that maternal verbal ability/knowledge could relate directly to cognitive test scores, and that race could relate directly to parenting. As such, we now additionally offer two alternative, *partial mediation* model specifications that we suspect might possibly, but not necessarily, play a role in explaining the race gap in cognitive test scores. These two alternative models (i.e., partial mediation models) are presented in Figure 1, Models B and C. Figure 1, Model B, specifies a direct effect from maternal verbal ability/knowledge to *g*, and we label this direct effect a *verbal socialization* effect. The notion of *verbal socialization* implies a direct impact of a mothers' language knowledge and usage on a child's knowledge of reading, vocabulary, and mathematics (Dollaghan et al., 1999; Schady, 2011). That is, because a mother's degree of verbal capability represents the complexity of the medium through which the child learns cognitive knowledge and skills at home, children of mothers with higher verbal ability should directly acquire more advanced cognitive skills and knowledge. Although this relationship is virtually self-evident in the case of vocabulary (i.e., children learn the vocabulary words their mothers use), we assert it is also true for children's development of reading comprehension and mathematics skills, which mothers often explain to children verbally.

Next, in Figure 1, Model C, we depict the second alternative theoretical model specification, in which race can have direct effects on parenting factors. We refer to these direct effects as *culturally specific parenting*. There exist a number of reasons why parenting factors might differ by race. First, the economic disadvantage faced by many African American parents has a debilitating effect on parents' ability to provide supportive, consistent, and involved parenting. Specifically, poor parents and African American parents have been found to be less responsive to a child's socioemotional needs, show fewer expressions of affection, and scold their children more frequently (Mills-Koonce et al., 2011; Peterson & Peters, 1985; Wilson, 1974). Low-income families often grow up in more dangerous neighborhoods, in which a child who disobeys his or her parents is at greater risk for involvement in various kinds of antisocial or dangerous activity (as victims of crime, as perpetrators, or as victims of police brutality; Kelley, Power, & Wimbush, 1992; Williams & Kornblum, 1985). Because African American families tend to be situated in more disadvantaged and dangerous neighborhoods (Barbarin & Soler, 1993; Krieger, Rowley, Herman, Avery, & Phillips, 1993), they are more likely to be targets of violence (e.g., Terrill & Reisig, 2003, showed that members of ethnic minority groups are more likely to be targets of police use of violent force, and this effect is explained/mediated by neighborhood disadvantage). As such, it is sometimes appropriate, adaptive, and caring for Black parents to use a stricter parenting style to protect children from being involved with dangerous activities and people (Kelley et al., 1992), as well as to socialize their children in how to respond quickly to authoritarian figures, such as police, who require compliance without offering explanations. These parenting practices can also lead to social norms, which would entail African American parents' using less permissive parenting styles, handed down from previous generations, regardless of current levels of SES. A more complete review of rationales for Black–White cultural differences in parenting (e.g., more use of punitive discipline and greater power assertion) is provided by McLoyd (1990). Ultimately, we hypothesize direct effects of race on parenting factors as an alternative model specification in the 3-Step Model (Figure 1, Model C).

Age as a Potential Boundary Condition for the 3-Step Model

Before we begin testing our model of the race gap in cognitive test scores, a reviewer asked us to use our longitudinal data to assess the extent to which our three-step explanatory model generalizes across developmental stages, from ages 4 to 15 years. For example, it is technically possible that some explanatory variables (e.g., birth order) might have a stronger effect on *g* at younger ages, but that the effect diminishes in magnitude at older ages. As such, we also plan to evaluate whether age is a boundary condition for our model.

Method

Participants and Procedure

Data for the current study were drawn from 1,364 families who participated in the National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development (SECCYD; 2005). Participants were recruited in or near 10 hospitals around the United States by the Early Child Care Research Network (ECCRN), beginning in 1991 up until 2009. Families were enrolled shortly after the child was born and were randomly selected from mothers giving birth during selected 24-hr sampling blocks. Cognitive tests were administered at five time points (54 months of age, first grade, third grade, fifth grade, and age 15 years). For various reasons, some families did not continue to participate throughout the entirety of the study. By Phase 4 of the study, when participants were in seventh through ninth grades (from 2005 to 2007), 1,009 families remained in the study. Given the focus of the current study (i.e., explaining the Black–White gap in cognitive test scores), we further restricted the sample to include only participants from Black and White demographic groups (i.e., we dropped 11% of the sample who were not identified as White/non-Hispanic or African American, so the final sample was 14% Black and 86% White). More information about the recruitment and selection procedures can be found in NICHD Early Child Care Research Network (2005) or online (see NICHD, 2012).

Due to missing data, different variables had different sample sizes (*N*s for cognitive tests ranged from 954 [Time 1] to 791 [Time 5]). To reduce missing data bias in the longitudinal model parameters, the current study's analyses were based upon a covariance matrix estimated via the expectation maximization (EM) algorithm, which is a maximum likelihood missing data technique that yields much less missing data bias than listwise or pairwise deletion strategies (Enders, 2010; Graham, 2009; Newman, 2003, 2009; Newman & Cottrell, 2015; Schafer & Graham, 2002). One of the reasons for the substantial missing data bias found under listwise deletion is that the data analyst is discarding real data provided by the partial respondents (Newman, 2014). In addition to introducing substantial bias, listwise deletion has the further disadvantage of reducing statistical power. Nonetheless, a reviewer asked us to check whether results of the current study might change if we had used listwise deletion instead of the EM algorithm. In deference to this reviewer, we reran all models using listwise deletion ($N = 484$), and results (estimates, significance tests, fit indices) remained essentially the same.

Measures

Cognitive ability. Cognitive ability was measured using the math, vocabulary, and reading ability facets of the Woodcock-Johnson Psycho-Educational Battery-Revised (WJ-R; Woodcock, 1990; Woodcock & Johnson, 1989). We chose to use these three ability facets because these were the only facets of ability measured in the current dataset for each child at all five points in time: 54 months of age, first grade, third grade, fifth grade, and 15 years of age. Math was measured using the Applied Problems subtest, which assesses the use of mathematical skills, such as adding or subtracting, to solve practical problems. Vocabulary was measured using the Picture-Vocabulary subtest, which assesses children's ability to identify pictured objects by name. Reading for the first four time points (through fifth grade) was measured with the Letter-Word Identification subtest, which assesses children's ability to identify printed letters and words, as well as their ability to match words to pictures. At 15 years of age, the Letter-Word Identification test was not administered; so the reading subtest was the Passage Comprehension subtest, which assesses children's ability to identify missing words in a passage and their ability to match word phrases to pictures. At each point in time, Math, Vocabulary, and Reading subtest scores were used together to reflect general cognitive ability/knowledge. Cognitive subtest scores were scaled as W scores (Jaffe, 2009; Woodcock & Dahl, 1971). The W scale is an equal-interval scale derived via item response theory, which places item difficulty and ability level on the same metric so that they are comparable. This allows researchers to examine growth in scores independent of the change in difficulty of the exam, and allows scores on the scale to be compared across ability levels and across ages.

Explanatory variables. Our analyses examine several explanatory variables that we expect to account for the relationship between race and cognitive test scores: birth order, maternal verbal ability/knowledge, learning materials, maternal sensitivity, maternal acceptance, physical environment, birth weight, and income.

Birth order. Birth order data were collected during the researchers' first visit to the family home, which took place when the child was 1 month old. A higher number indicates that the child was born later ($1 =$ firstborn, $2 =$ secondborn, etc.). This variable ranged from 1 to 7.

Maternal verbal ability/knowledge. Maternal verbal ability/knowledge was measured using the Peabody Picture Vocabulary Test–Revised. This test was administered when the child was 36 months old. Split-half reliabilities for this measure ranged between .80 and .83.

Maternal sensitivity. Maternal sensitivity ratings were obtained via a videotaped interaction between a mother (or in rare cases [less than 5%], another primary caregiver such as a father or grandparent) and her child, in a laboratory. Children were each asked to complete a set of semistructured tasks with the help of their mother, and maternal sensitivity was coded by multiple raters using 7-point rating scales (all intercoder reliabilities $> .80$). Observations lasted between 15 and 20 min, during which the mother was judged on responses to her child's needs (e.g., distress) and the extent to which she was engaged with her child in using the toys provided (e.g., rattle, balls for 6 months old). An example of a specific task is using an Etch-A-Sketch to traverse a maze. Parents were rated on a composite of supportive presence, respect for autonomy of the child, and reflected hostility (reverse coded). Data for maternal sensitivity used in this study were collected at several time points leading up to 54 months of age: 6 months ($\alpha = .75$), 15 months ($\alpha = .70$), 24 months ($\alpha = .79$), 36 months ($\alpha = .78$), and 54 months ($\alpha = .84$).

Learning materials. For three variables (learning materials, maternal acceptance, and physical environment), trained interviewers asked parents a series of questions in a specified order while allowing for changes if needed (e.g., more warm-up questions to establish rapport between interviewer and mother). In addition, interviewers used their own observations to answer certain questions (e.g., "Outside play environment appears safe," "Parent does not scold or derogate child more than once"). All interview questions were scored in a binary fashion (1 = yes, 0 = no), and items were summed to form subscale composite scores.

Information about stimulation in, and quality of, the home environment was collected using HOME (Caldwell & Bradley, 1984). This measure was administered to mothers as a semistructured interview when the child was 54 months of age. Questions in this interview focused on the types of family experiences, both in and out of the home, that have been theorized to foster social and cognitive development (Bradley et al., 1989). The Learning Materials subscale, an 11-item measure, assesses the extent to which the child has access to learning materials (e.g., "Child has toys which teach color, size, and shapes," "Child has three or more puzzles," "Child has at least 10 children's books"). The internal consistency reliability was $\alpha = .57$ for the Learning Materials subscale.

Maternal warmth and acceptance. The Acceptance subscale of the HOME was administered to mothers using a semistructured interview at 54 months. This subscale measures how the mother interacts with the child, and the extent to which the mother accepts imperfect behavior and avoids punishing the child harshly (e.g., "Parent does not scold or derogate child more than once," "No more than one instance of physical punishment during last week"). The internal consistency reliability was $\alpha = .52$ for this four-item subscale.

Physical environment. The seven-item Physical Environment subscale (measured as part of the HOME semistructured interview at 54 months) assesses the extent to which parents provide a safe and clean home environment (e.g., "Building appears safe and free from hazards," "Rooms are not overcrowded with furniture," "House is reasonably clean and minimally cluttered"). The internal consistency reliability was $\alpha = .63$ for the Physical Environment subscale. Although several of our predictors were found to have low internal consistency reliabilities, these measures were shown to be valid predictors of cognitive development in both the current and past studies (maternal warmth and acceptance, Brooks-Gunn et al., 2003; physical environment, Bradley et al., 1988). Further, if the measures in question had been more reliable, then the observed relationships would have likely become stronger (disattenuated), not weaker.

Birth weight. Child birth weight was reported by the mother. Data were collected by research associates in the hospital on the day the child was born. Birth weight data were coded in 100-g increments.

SES/Income-to-needs. We used family income-to-needs and maternal education as our indicators of SES. Income-to-needs was measured at 6 months, 15 months, 24 months, 36 months, and 54 months of age, and was a computed as a ratio of income-to-needs. This ratio is made up of the total household income divided by the federal index for poverty for a given family size. For example, the poverty line in the year 2012 for a family of two parents and two children was $23,283 (United States Census Bureau, 2012b). Maternal education was measured when the child was 1 month old as part of a set of demographic variables. This variable was measured by the number of years of schooling mothers said they had completed (e.g., 12 being a high school diploma, 14 being some college).

Analyses

Analyses were conducted in two parts, with all model specification details given in Appendix B. First, we estimated a sequence of a priori models (Models 1a, 1b, 2, and 3; described in Appendix B), to establish measurement equivalence of the cognitive tests over time and to demonstrate that the covariates described previously can explain the race gap in *g*. Second, we estimated the three, more parsimonious theoretical models shown in Figure 1 (Model A: 3-Step Model; Model B: 3-Step Model with Verbal Socialization; and Model C: 3-Step Model with Verbal Socialization and Culturally Specific Parenting). The 3-Step Models have the same model specifications as Model 3 in Appendix B, except that the 3-Step Models specify a particular pattern of relations among the explanatory variables.

Results

Table 2 presents the correlation matrix among constructs in this study, as well as means and standard deviations. Large race gaps in cognitive ability test scores are present at every time point (*d*s range from -1.24 to -1.39, $p < .05$). As described in Appendix B, measurement equivalence was established for the cognitive tests over time, and race was strongly related to the *g* intercept ($\gamma = -0.42$, $p < .05$, $d = -1.35$), but race was not related to the *g* slope ($\gamma = -0.05$, $p > .05$, *ns*, $d = -.11$; i.e., from Model 2, which is found in Appendix B). Further, after adding the explanatory variables to the model, race no longer had a statistically significant direct relation to the *g* intercept ($\beta = -0.07$, $p > .05$, *ns*; see Table B1, Appendix B, Model 3). Table B1 also displays the percent of the race gap that can be explained by each of the explanatory variables (described in Appendix B). We strongly recommend reading Appendix B and Table B1.

3-Step Model of the Cognitive Test Score Gap

The 3-Step Model (i.e., full mediation model, Figure 2a) shows adequate fit to our data Comparative Fit Index [CFI] = .95, Non-

Table 2
Correlation Matrix Among Latent Variables

	M	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. *g* time 1	456.78	11.92	—														
2. *g* time 2	471.06	11.35	.92*	—													
3. *g* time 3	485.35	11.23	.91*	.95*	—												
4. *g* time 4	499.63	11.07	.90*	.96*	.97*	—											
5. *g* time 5	535.34	11.84	.81*	.87*	.91*	.95*	—										
6. *g* intercept	477.16	11.92	.94*	.97*	.97*	.97*	.88*	—									
7. Birth order	1.82	0.94	−.24*	−.25*	−.25*	−.25*	−.23*	−.26*	—								
8. Maternal verbal ability/knowledge	99.01	18.11	.57*	.61*	.62*	.66*	.61*	.62*	−.11*	—							
9. Learning materials	9.39	1.51	.44*	.46*	.47*	.48*	.45*	.47*	−.11*	.45*	—						
10. Maternal sensitivity	9.24	1.24	.59*	.61*	.62*	.62*	.57*	.63*	−.06	.62*	.55*	—					
11. Maternal acceptance	4.38	0.80	.36*	.38*	.38*	.37*	.34*	.38*	−.03	.36*	.37*	.49*	—				
12. Physical environment	6.31	1.11	.39*	.40*	.40*	.40*	.35*	.41*	−.12*	.31*	.37*	.49*	.35*	—			
13. Birth weight	34.90	5.05	.15*	.15*	.16*	.16*	.14*	.16*	.04	.16*	.07	.18*	.07	.10*	—		
14. Income-to-needs	3.65	2.71	.43*	.45*	.44*	.45*	.39*	.46*	−.21*	.46*	.44*	.51*	.28*	.35*	.04	—	
15. Maternal education	14.31	2.49	.48*	.51*	.53*	.54*	.50*	.52*	−.11*	.63*	.50*	.61*	.37*	.35*	.11*	.58*	
16. Race (*r*) (W = 0, B = 1)	0.14	0.35	−.40*	−.42*	−.43*	−.44*	−.42*	−.43*	.13*	−.43*	−.35*	−.51*	−.26*	−.31*	−.19*	−.30*	−.23*
B Mean (*SD*)			443.36	462.22	476.16	490.31	524.85	468.60	2.12	80.29	8.19	7.91	3.88	5.48	32.61	1.86	12.89
			(12.21)	(9.14)	(10.79)	(10.79)	(11.93)	(9.69)	(1.20)	(15.56)	(1.75)	(0.89)	(0.90)	(1.49)	(4.09)	(1.15)	(1.63)
W Mean (*SD*)			459.04	472.55	486.90	501.20	537.11	478.61	1.78	102.23	9.61	9.47	4.47	6.45	35.29	3.96	14.55
			(14.69)	(8.94)	(9.02)	(8.91)	(10.68)	(8.46)	(0.87)	(16.42)	(1.37)	(0.92)	(0.76)	(0.97)	(5.09)	(2.75)	(2.53)
Race (*d*)			−1.24*	−1.32*	−1.35*	−1.39*	−1.32*	−1.35*	.37*	−1.37*	−1.05*	−1.70*	−.78*	−.91*	−.54*	−.88*	−.69*

Note. Race subgroup *d*s are approximated from race *r*s, using the formula $d = \frac{r}{\sqrt{(1-r^2)(p(1-p))}}$ (Lipsey & Wilson, 2000). •••.

* $p < .05$.

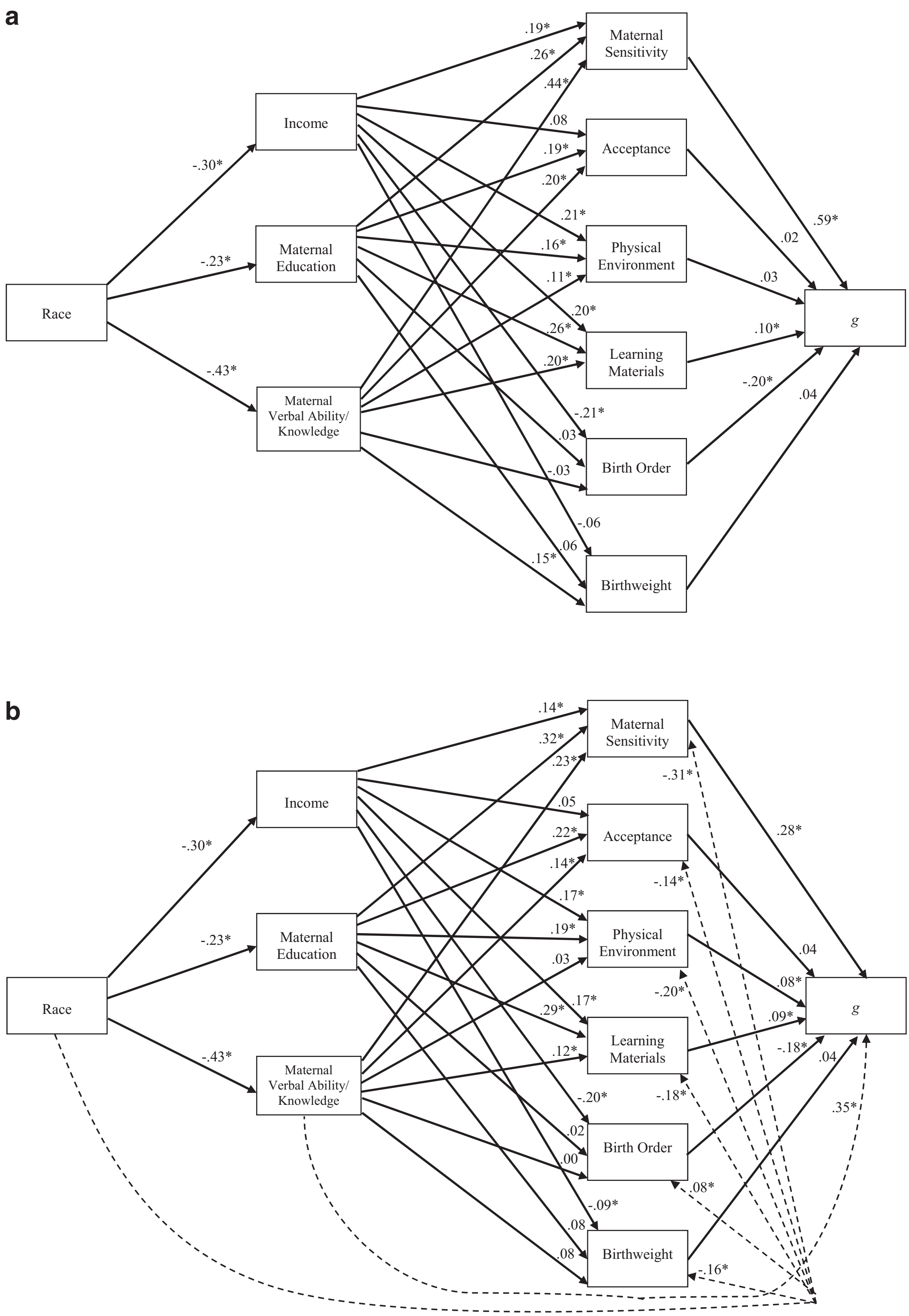

Figure 2. (a) 3-Step Model of race and cognitive test scores (full mediation; coefficients are standardized). (b) 3-Step Model plus verbal socialization and culturally specific parenting (coefficients are standardized).

Table 3
Summary of 3-Step Model Fit

Model	$\chi^2(df)$	CFI	NNFI	RMSEA [90% CI]
Model A: 3-Step Model	1384.30 (441)	.947	.937	.052 [.049, .055]
Model B: 3-Step Model with Verbal Socialization	1308.80 (440)	.951	.942	.050 [.047, .053]
Model C: 3-Step Model with Verbal Socialization and Culturally Specific Parenting	1191.80 (434)	.958	.948	.047 [.044, .050]

Note. *df* = degrees of freedom; CFI = comparative fit index; NNFI = non-normed fit index; RMSEA = root mean square error of approximation; CI = confidence interval.

Normed Fit Index [NNFI] = .94, Root Mean Squared Error of Approximation [RMSEA] = .052, and R_g^2 = .52). As shown in Figure 2a, race is significantly related to all three maternal advantage factors: income ($\beta = -.30$, $SE = .03$, $p < .05$), maternal education ($\beta = -.23$, $SE = .03$, $p < .05$), and maternal verbal ability/knowledge ($\beta = -.43$, $SE = .03$, $p < .05$). Maternal advantage, in turn, leads to the acquisition of parenting factors, including maternal sensitivity, acceptance, physical environment, and learning materials (11 of 12 parameter estimates from maternal advantage to these parenting factors are statistically significant; $p < .05$). Birth order is only predicted by income ($\beta = -.21$, $SE = .05$, $p < .05$), and birth weight is only predicted by maternal verbal ability/knowledge ($\beta = .15$, $SE = .05$, $p < .05$). Finally, maternal sensitivity ($\beta = .59$, $SE = .05$, $p < .05$), learning materials ($\beta = .10$, $SE = .04$, $p < .05$), and birth order ($\beta = -.20$, $SE = .04$, $p < .05$) are uniquely statistically related to *g*, whereas acceptance ($\beta = .02$, $SE = .04$, *ns*), physical environment ($\beta = .03$, $SE = .04$, *ns*), and birth weight ($\beta = .04$, $SE = .03$, *ns*) are not.

Figure 2b adds the two additional sets of paths specified previously: verbal socialization (from Model B in Table 3: CFI = .95, NNFI = .94, RMSEA = .050, and R_g^2 = .53) and culturally specific parenting (from Model C in Table 3: CFI = .96, NNFI = .95, RMSEA = .047, and R_g^2 = .54). The fit for Model C to our data is good. As predicted by the model, the verbal socialization parameter (*g* regressed on maternal verbal ability/knowledge) is statistically significant ($\beta = .35$, $SE = .04$, $p < .05$). Additionally, race was directly related to all of the parenting factors: maternal sensitivity ($\beta = -.31$, $SE = .03$, $p < .05$), acceptance ($\beta = -.14$, $SE = .04$, $p < .05$), physical environment ($\beta = -.20$, $SE = .04$, $p < .05$), learning materials ($\beta = -.18$, $SE = .03$, $p < .05$), birth order ($\beta = .08$, $SE = .04$, $p < .05$), and birth weight ($\beta = -.16$, $SE = .04$, $p < .05$). As shown in Figure 2b, all regression coefficients have the same statistical significance pattern and are in the same directions as in Figure 2a, with three exceptions: maternal verbal ability/knowledge is no longer statistically related to physical environment ($\beta = .03$, $SE = .05$, *ns*), or to birth weight ($\beta = .08$, $SE = .05$, *ns*), and income is significantly related to birth weight ($\beta = -.09$, $SE = .05$, $p < .05$).[1] Further, if a direct path from race to *g* were added to Model C (Figure 2b), it would not be statistically significant ($\beta = -.065$, $SE = .033$, *ns*). Finally, we note that whereas the practical fit indices for Models A, B, and C are fairly equivalent (see Table 3), the 3-Step Model with Verbal Socialization and Culturally Specific Parenting (Model C, Figure 2b) appears to exhibit slightly better fit with the present data, $\Delta\chi^2(6) = 117$; $p < .05$.

Age as a Potential Boundary Condition for the Model

On a different note, we can also address the possibility that child age might serve as a boundary condition for our 3-Step Model of the race gap in *g*. In order to test this possibility, we estimated the 3-Step Model shown in Figure 2b five separate times—once for the measure of *g* at each age. Model fit indices suggested adequate—and similar—fit for the 3-Step Model at each time point (CFI/NNFI/RMSEA ranged from .94/.93/.058 to .96/.94/.054). Further, the percent of the race gap explained by the model fell only modestly from 89% (at 54 months) to 85% (at first grade) to 79–80% (at third grade, fifth grade, and 15 years). All model parameter estimates were consistent over time, although the path from maternal verbal ability to *g* was slightly smaller at 54 months ($\gamma = .22$), and the path from maternal sensitivity to *g* was slightly larger at 54 months ($\gamma = .40$; cf. Figure 2b; these paths were consistently positive and statistically significant at each time point). In general, the 3-Step Model replicated well across ages, with the possibility that maternal sensitivity might slightly decrease in importance, and maternal verbal ability/knowledge might slightly increase in importance, with the onset of schooling. To avoid the possibility of overfitting the data by offering conclusions based merely on visual inspection of post hoc trends, we recommend that future research be conducted before drawing firm conclusions about changes in the explanatory power of each mechanism across ages.

Discussion

The purpose of the current article was to explain the origins of adverse impact. We did this by modeling Black–White cognitive test score gaps longitudinally between 54 months and 15 years of age (i.e., across the majority of the life course before individuals enter the workforce), and by attempting to offer an integrated theoretical model to explain this gap. We quantified the size of the gap over time, and also investigated the extent to which our explanatory variables (income, maternal education, maternal verbal scores, learning materials, maternal sensitivity, maternal acceptance, physical environment, birth order, and birth weight) could account for the association between race and cognitive test scores. Ultimately, we proposed a 3-Step Model that presents a more parsimonious explanation for how cognitive test gaps develop.

Our results suggest that Black–White gaps in cognitive test scores are large and pervasive, and are already established by 54 months of age. This is indicated by the mean differences in cognitive test scores at each time point (i.e., subgroup *d*s range from -1.24 to -1.39 across the time points; see Table 2), as well as the race gap in the *g* intercept ($d = -1.35$). Further, between 54

[1] Parameter estimates in the 3-Step Model with Verbal Socialization (not pictured), are all in the same direction and show the same pattern of statistical significance as in Figure 2a, with the single exception that the path from physical environment to *g* increases slightly (β increases from $\beta = .03$ [$SE = .04$], *ns*, to $\beta = .08$ [$SE = .03$], $p < .05$) when the verbal socialization path is added.

months and 15 years of age, this gap did not significantly increase over time, as indicated by the lack of relation between race and the *g* slope from the linear growth model.

Figures 2a and 2b present our 3-Step Models. The 3-Step Model examines the process by which race leads to cognitive test score gaps. In particular, we show that race gives rise to a set of group differences in maternal advantage factors: income, maternal education and maternal verbal ability/knowledge (Step 1), which in turn lead to parenting factors of maternal sensitivity, acceptance, physical environment, learning materials, birth weight, and birth order (Step 2), which in turn promote cognitive ability/knowledge in children (Step 3). We also found that maternal verbal ability/knowledge directly impacts a child's cognitive ability/knowledge, a phenomenon we refer to as verbal socialization. Finally, we examined differences in parenting by race, a phenomenon we refer to as culturally specific parenting. The 3-Step Model with Verbal Socialization and Culturally Specific Parenting was most strongly supported in our data.

The implication of this research is that adverse impact created by cognitive tests may arise from Black–White differences in the important developmental conditions described above. The current theoretical model therefore contributes to theories about the origins of subgroup differences in cognitive test scores, which has been cited as a major theoretical gap in current models of adverse impact (Outtz, 2010). Interestingly, all of the explanatory covariates in the current research were measured at or before age 54 months—prior to the onset of formal schooling. Although this result does *not* mean that schooling is an unimportant factor in increasing or decreasing race gaps in cognitive achievement, the result does suggest that the vast majority (over 80% in the current study; see Table B1) of the Black–White achievement gap can be descriptively explained using prospective factors that occurred prior to the onset of formal schooling. This result at least implies that interventions targeted at home environment, maternal sensitivity, learning materials, and birth order/household size might be worth studying as means to interrupt the cycle of disadvantaged cognitive test performance. Because the current study was nonexperimental, future experimental work will be needed before concrete recommendations can be made.

Understanding how and why adverse impact develops is a critical first step toward discerning what scientists and practitioners can do to address, and perhaps prevent, such gaps in the future. Many of the factors that contribute to race differences in cognitive test scores might best be addressed by specialists in the fields of education, public policy, or child psychology, whereas some other factors contributing to the gap (e.g., group differences in income due to access to occupations obtained via hiring and admissions procedures) might be more immediately addressed by I/O-Human Resources (HR) practitioners. Given the size of the gap, our own speculation is that removing the race gap in test scores will require effort from a diverse group of specialists working simultaneously, and will be accelerated if specialists from diverse disciplines each take responsibility for those aspects of the problem to which they can contribute solutions.

Results of this article also suggest new directions for adverse impact research. Namely, researchers should continue to examine the extent to which the different societal and developmental resources that create cognitive test score gaps (Table B1) might also create gaps in actual job performance. This is an essential question for personnel selection scientists and practitioners, given that race gaps in cognitive tests are approximately 3 times larger than corresponding race gaps in job performance (McKay & McDaniel, 2006; Outtz & Newman, 2010). Such studies have the potential to develop a fuller picture of which racial inequalities (or inequities) must be addressed in order to reduce adverse impact in hiring and admissions.

To elaborate, some models of test fairness suggest that the key problem of adverse impact is due to elements of cognitive tests that overlap with race but do not overlap with job performance (Cole, 1973; Darlington, 1971; Newman et al., 2007). Outtz and Newman (2010) refer to this as *performance-irrelevant race-related variance* in cognitive test scores. If this aspect of cognitive test scores is large, it implies that when cognitive tests are used for hiring, African Americans would tend to be excluded from jobs for reasons that have nothing to do with job performance. Because the current study does not include any measures of job performance (i.e., the sample was not old enough to be legally employed), we cannot presently address the development of performance-irrelevant race-related cognitive test score variance. That is, some of the racially disparate conditions that give rise to cognitive test score gaps might also give rise to similar Black–White job performance gaps, whereas other racially disparate conditions that explain cognitive test score gaps might not give rise to similar job performance gaps. Teasing apart these conditions that have larger or smaller job performance implications will be an important next step (e.g., does the unavailability of learning materials in the home harm cognitive test performance much more than it harms actual job performance?).

Another potential direction for future research is to change cognitive tests themselves so that they retain their high validity while reducing adverse impact. This could involve changing the way test material is presented (Schmitt & Quinn, 2010) as well as exploring the extent to which cognitive test questions may be race-loaded. For example, technical knowledge tests tend to show much larger Black–White differences than do math tests or cognitive speed tests (Alderton, Wolfe, & Larson, 1997; Hough, Oswald, & Ployhart, 2001; Kehoe, 2002; Outtz & Newman, 2010; Wee et al., 2014). This might be because measures of some facets of cognitive ability/knowledge also unintentionally assess aspects of socially privileged life experience, including familiarity with testing styles and situations (H. W. Goldstein et al., 2010). Thus, one potential way to reduce Black–White cognitive test gaps is to create a cognitive test that is unfamiliar to all participants while still being a valid measure of cognitive ability/knowledge. Although this might not eliminate adverse impact altogether, such a strategy could eliminate contamination of the cognitive test due to privileged experience (H. W. Goldstein et al., 2010; Yusko & Goldstein, 2008). A revival of research on the construct of intelligence might help solve these and other fundamental questions regarding the use of cognitive tests in hiring and admissions decisions (see Scherbaum, Goldstein, Yusko, Ryan, & Hanges, 2012).[2]

Limitations

This article has several limitations. One limitation is that the NICHD SECCYD is not a strictly random probability sample of

[2] We note that race gaps in cognitive test scores do not imply that cognitive tests suffer from measurement bias (Drasgow, 1984), nor do racial group differences on tests necessarily imply that cognitive tests lack construct validity. Instead, the current research offers a theoretical reason why group differences in cognitive tests would be observed, even in the absence of test bias. The current article shows that the gap can be explained via sociological and developmental mechanisms, without invoking psychometric explanations.

the U.S. population. Families were not eligible for the SECCYD if the mother was under 18 years of age, did not speak English, or had a substance abuse or other serious health problem. Additionally, if the child was hospitalized for more than 7 days after birth, had disabilities, had a twin, or if the family was in a neighborhood that was too dangerous or too far from the study site, they were not eligible to participate.

Next, it is imperative for readers to remember that the current study is based on correlational data only. As such, our model represents only one sufficient explanation that is consistent with these data, and does not support causal inferences involving the explanatory mechanisms. There may exist other possible explanations for the test score gap that were not included in past research and therefore were not included here.

In addition, our research has focused on mothers' attributes, largely ignoring fathers. A reviewer raised the question of whether children in father-absent families tend to have lower cognitive test scores than children in two-parent homes, and indeed research supports this idea (Santrock, 1972; Shinn, 1978). Nonetheless, research also shows that the effect of father absence on cognitive development is no longer significant after controlling for other factors, especially SES (e.g., H. S. Goldstein, 1982; Svanum, Bringle, & McLaughlin, 1982). This suggests father absence is causally upstream from—and operates by way of—SES, but has no unique power in explaining the achievement gap. As a supplemental analysis, we used the SECCYD data to create a composite measure of whether the father has lived in the home, averaging across points in time from birth through age 15 ($\alpha = .98$), and this father-absence composite was unrelated to *g* after controlling for the other variables in our explanatory model ($t = 0.17$; $p > .05$). In other words, father absence adds no explanatory value beyond the variables already in our model.

Another important point, also noted by a reviewer, is that the current study modeled cognitive change/slope differences over time by using physical age to represent time. Besides *age*, the notion of "time" might alternatively correspond to changes that transpired during a particular historical *period* or across birth *cohorts* (see Schaie, 1965). As such, the current article provides only one explanation for the manner in which cognitive test gaps might emerge and change (i.e., across *age*). These findings serve as a starting point for additional research on possible changes in test gaps across birth cohorts and across dates/historical periods.

Both the full model of this study (Model 3, Table B1) and the 3-Step Model with Verbal Socialization and Culturally Specific Parenting (Figure 2b) have explained the Black–White cognitive test score gap (i.e., the coefficient of race directly predicting *g* was no longer statistically significant). However, we acknowledge that the direct relation was not precisely zero ($\beta = .07$, $p > .05$, *ns*) and is similar in size to several of the indirect effects found in our analysis. Future studies should examine other possible explanatory variables for the Black–White cognitive test score gap in conjunction with the variables already present in our model.

Finally, the current data set does not allow us to investigate potentially interesting research questions brought up in previous research, such as the effects of summer learning versus school learning over time (e.g., Alexander, Entwisle, & Olson, 2007; Downey, von Hippel, & Broh, 2004). Further, there is no publicly available data on the cognitive development of these participants beyond 15 years of age. Future studies should examine Black–White cognitive test score gaps as individuals continue into the workforce, to assess the possibility that work experience might enhance or ameliorate the cognitive gap for individuals in certain occupations.

Conclusion

We examined the extent to which specific developmental conditions could account for Black–White gaps in cognitive test scores, to theoretically explain the origins of adverse impact. Results showed that Black–White gaps were large at every time point from 54 months to 15 years of age, but that the gap did not grow (or shrink) over time. Finally, we accounted for the association between race and cognitive test scores using our covariates, and provided a parsimonious theoretical model that explicates relations among the covariates, offering a 3-Step Model of the relations among the racially disparate conditions that appear to give rise to the race gap in test scores. This study therefore attempts to fill a hole in adverse impact theory by pinpointing how cognitive test score gaps might arise.

References

Ainsworth, M. D., Bell, S. M., & Stayton, D. J. (1974). Infant–mother attachment and social development: Socialization as a product of reciprocal responsiveness to signals. In M. P. Richards (Ed.), *The Integration of a Child into a Social World* (pp. 99–135). New York, NY: Cambridge University Press.

Alderton, D. L., Wolfe, J. H., & Larson, G. E. (1997). The ECAT battery. *Military Psychology, 9,* 5–37. http://dx.doi.org/10.1207/s15327876mp0901_1

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007). Lasting consequences of the summer learning gap. *American Sociological Review, 72,* 167–180. http://dx.doi.org/10.1177/000312240707200202

Bacharach, V. R., & Baumeister, A. A. (1998). Effects of maternal intelligence, marital status, income, and home environment on cognitive development of low birthweight infants. *Journal of Pediatric Psychology, 23,* 197–205. http://dx.doi.org/10.1093/jpepsy/23.3.197

Barbarin, O. A., & Soler, R. E. (1993). Behavioral, emotional, and academic adjustment in national probability sample of African-American children: Effects of age, gender, and family structure. *Journal of Black Psychology, 19,* 423–446. http://dx.doi.org/10.1177/00957984930194004

Bargh, J. A., & Schul, Y. (1980). On the cognitive benefit of teaching. *Journal of Educational Psychology, 72,* 593–604. http://dx.doi.org/10.1037/0022-0663.72.5.593

Becker, S., & Lewis, H. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy, 81,* S279–S288. http://dx.doi.org/10.1086/260166

Behrman, J. R., & Taubman, P. (1986). Birth order, schooling, and earnings. *Journal of Labor Economics, 4,* S121–S150. http://dx.doi.org/10.1086/298124

Bennett, D. S., Bendersky, M., & Lewis, M. (2008). Children's cognitive ability from 4 to 9 years old as a function of prenatal cocaine exposure, environmental risk, and maternal verbal intelligence. *Developmental Psychology, 44,* 919–928. http://dx.doi.org/10.1037/0012-1649.44.4.919

Black, S. E., Devereux, P. G., & Salvanes, K. G. (2005). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics, 120,* 669–700.

Bobko, P., & Roth, P. L. (2013). Reviewing, categorizing, and analyzing the literature on Black White mean differences for predictors of job performance: Verifying some perceptions and updating/correcting others. *Personnel Psychology, 66,* 91–126. http://dx.doi.org/10.1111/peps.12007

Bobko, P., Roth, P. L., & Potosky, D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology, 52,* 561–589. http://dx.doi.org/10.1111/j.1744-6570.1999.tb00172.x

Booth, A. L., & Kee, H. J. (2009). Birth order matters: The effect of family size and birth order on educational attainment. *Journal of Population Economics, 22,* 367–397. http://dx.doi.org/10.1007/s00148-007-0181-4

Bradley, R. H., & Caldwell, B. M. (1984). The HOME inventory and family demographics. *Developmental Psychology, 20,* 315–320. http://dx.doi.org/10.1037/0012-1649.20.2.315

Bradley, R. H., Caldwell, B. M., Brisby, J., Magee, M., Whiteside, L., & Rock, S. L. (1992). The HOME inventory: A new scale for families of pre- and early adolescent children with disabilities. *Research in Developmental Disabilities, 13,* 313–333. http://dx.doi.org/10.1016/0891-4222(92)90009-U

Bradley, R. H., Caldwell, B. M., Rock, S. L., Hamrick, H. M., & Harris, P. (1988). Home observation for measurement of the environment: Development of a home inventory for use with families having children 6 to 10 years old. *Contemporary Educational Psychology, 13,* 58–71. http://dx.doi.org/10.1016/0361-476X(88)90006-9

Bradley, R. H., Caldwell, B. M., Rock, S. L., Ramey, C. T., Barnard, K. E., Gray, C., . . . Johnson, D. L. (1989). Home environment and cognitive development in the first 3 years of life: A collaborative study involving six sites and three ethnic groups in North America. *Developmental Psychology, 25,* 217–235. http://dx.doi.org/10.1037/0012-1649.25.2.217

Bradley, R. H., Corwyn, R. F., Burchinal, M., McAdoo, H. P., & Coll, C. G. (2001). The home environments of children in the United States part II: Relations with behavioral development through age thirteen. *Child Development, 72,* 1868–1886. http://dx.doi.org/10.1111/1467-8624.t01-1-00383

Bradley, R. H., Corwyn, R. F., McAdoo, H. P., & Coll, C. G. (2001). The home environments of children in the United States part I: Variations by age, ethnicity, and poverty status. *Child Development, 72,* 1844–1867. http://dx.doi.org/10.1111/1467-8624.t01-1-00382

Brooks-Gunn, J., & Duncan, G. J. (1997). The effects of poverty on children. *The Future of Children, 7,* 55–71. http://dx.doi.org/10.2307/1602387

Brooks-Gunn, J., Klebanov, P. K., Smith, J., Duncan, G. J., & Lee, K. (2003). The Black–White test score gap in young children: Contributions of test and family characteristics. *Applied Developmental Science, 7,* 239–252. http://dx.doi.org/10.1207/S1532480XADS0704_3

Brooks-Gunn, J., & Markman, L. B. (2005). The contribution of parenting to ethnic and racial gaps in school readiness. *The Future of Children, 15,* 139–168. http://dx.doi.org/10.1353/foc.2005.0001

Burchinal, M., McCartney, K., Steinberg, L., Crosnoe, R., Friedman, S. L., McLoyd, V., & Pianta, R., & the NICHD Early Child Care Research Network. (2011). Examining the Black–White achievement gap among low-income children using the NICHD study of early child care and youth development. *Child Development, 82,* 1404–1420. http://dx.doi.org/10.1111/j.1467-8624.2011.01620.x

Caldwell, B., & Bradley, R. (1984). *Administration manual: Home Observation for Measurement of the Environment*. Little Rock, AR: University of Arkansas.

Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). *Organizational Research Methods, 1,* 421–483. http://dx.doi.org/10.1177/109442819814004

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9,* 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5

Cole, N. S. (1973). Bias in selection. *Journal of Educational Measurement, 10,* 237–255. http://dx.doi.org/10.1111/j.1745-3984.1973.tb00802.x

Condron, D. J. (2009). Social class, school and non-school environments, and Black/White inequalities in children's learning. *American Sociological Review, 74,* 685–708. http://dx.doi.org/10.1177/000312240907400501

Crosnoe, R. (2009). Low-income students and the socioeconomic composition of public high schools. *American Sociological Review, 74,* 709–730. http://dx.doi.org/10.1177/000312240907400502

Cutler, D. M., Glaeser, E. L., & Vigdor, J. L. (1999). The rise and decline of the American ghetto. *Journal of Political Economy, 107,* 455–506. http://dx.doi.org/10.1086/250069

Darlington, R. B. (1971). Another look at "cultural fairness." *Journal of Educational Measurement, 8,* 71–82. http://dx.doi.org/10.1111/j.1745-3984.1971.tb00908.x

De Corte, W., Lievens, F., & Sackett, P. R. (2007). Combining predictors to achieve optimal trade-offs between selection quality and adverse impact. *Journal of Applied Psychology, 92,* 1380–1393. http://dx.doi.org/10.1037/0021-9010.92.5.1380

De Corte, W., Sackett, P., & Lievens, F. (2010). Selecting predictor subsets: Considering validity and adverse impact. *International Journal of Selection and Assessment, 18,* 260–270. http://dx.doi.org/10.1111/j.1468-2389.2010.00509.x

de Kieviet, J. F., Zoetebier, L., van Elburg, R. M., Vermeulen, R. J., & Oosterlaan, J. (2012). Brain development of very preterm and very low-birthweight children in childhood and adolescence: A meta-analysis. *Developmental Medicine & Child Neurology, 54,* 313–323. http://dx.doi.org/10.1111/j.1469-8749.2011.04216.x

De Wolff, M. S., & van IJzendoorn, M. H. (1997). Sensitivity and attachment: A meta-analysis on parental antecedents of infant attachment. *Child Development, 68,* 571–591. http://dx.doi.org/10.1111/j.1467-8624.1997.tb04218.x

Dezoete, J. A., MacArthur, B. A., & Tuck, B. (2003). Prediction of Bayley and Stanford-Binet scores with a group of very low birthweight children. *Child: Care, Health and Development, 29,* 367–372. http://dx.doi.org/10.1046/j.1365-2214.2003.00349.x

Dollaghan, C. A., Campbell, T. F., Paradise, J. L., Feldman, H. M., Janosky, J. E., Pitcairn, D. N., & Kurs-Lasky, M. (1999). Maternal education and measures of early speech and language. *Journal of Speech, Language, and Hearing Research, 42,* 1432–1443. http://dx.doi.org/10.1044/jslhr.4206.1432

Dotterer, A. M., Iruka, I. U., & Pungello, E. (2012). Parenting, race, and socioeconomic status: Links to school readiness. *Family Relations: An Interdisciplinary Journal of Applied Family Studies, 61,* 657–670. http://dx.doi.org/10.1111/j.1741-3729.2012.00716.x

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. http://dx.doi.org/10.1177/000312240406900501

Drasgow, F. (1984). Scrutinizing psychological tests: Measurement equivalence and equivalent relations with external variables are the central issues. *Psychological Bulletin, 95,* 134–135. http://dx.doi.org/10.1037/0033-2909.95.1.134

Duncan, G. J., & Magnuson, K. A. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps? *The Future of Children, 15,* 35–54. http://dx.doi.org/10.1353/foc.2005.0004

Dunn, L., & Dunn, L. (1981). *The Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service.

Duyme, M., Dumaret, A. C., & Tomkiewicz, S. (1999). How can we boost IQs of "dull" children? A late adoption study. *Proceedings of the National Academy of Sciences, USA, 96,* 8790.

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.

Equal Employment Opportunity Commission. (1978). *Uniform guidelines on employee selection procedures*. Washington, DC: Author.

Estrada, P., Arsenio, W. F., Hess, R. D., & Holloway, S. D. (1987). Affective quality of the mother-child relationship: Longitudinal consequences for children's school-relevant cognitive functioning. *Developmental Psychology, 23,* 210–215. http://dx.doi.org/10.1037/0012-1649.23.2.210

Feagin, J. R. (2006). *Systemic racism: A theory of oppression*. New York, NY: Routledge.

Fox, R. A., Platz, D. L., & Bentley, K. S. (1995). Maternal factors related to parenting practices, developmental expectations, and perceptions of child behavior problems. *The Journal of Genetic Psychology: Research and Theory on Human Development, 156,* 431–441. http://dx.doi.org/10.1080/00221325.1995.9914835

Fraley, R. C., Roisman, G. I., & Haltigan, J. D. (2013). The legacy of early experiences in development: Formalizing alternative models of how early experiences are carried forward over time. *Developmental Psychology, 49,* 109–126. http://dx.doi.org/10.1037/a0027852

Fryer, R. G., & Levitt, S. D. (2004). Understanding the Black–White test score gap in the first two years of school. *The Review of Economics and Statistics, 68,* 551–560.

Fryer, R. G., & Levitt, S. D. (2006). The Black–White test score gap through third grade. *American Law and Economics Review, 8,* 249–281. http://dx.doi.org/10.1093/aler/ahl003

Garcia Coll, C., Lamberty, G., Jenkins, R., McAdoo, H. P., Crnic, K., Wasik, B. H., & Garcia, H. V. (1996). An integrative model for the study of developmental competencies in minority children. *Child Development, 67,* 1891–1914. http://dx.doi.org/10.2307/1131600

Goldstein, H. S. (1982). Fathers' absence and cognitive development of 12- to 17-year-olds. *Psychological Reports, 51,* 843–848. http://dx.doi.org/10.2466/pr0.1982.51.3.843

Goldstein, H. W., Scherbaum, C. A., & Yusko, K. P. (2010). Revisiting g: Intelligence, adverse impact, and personnel selection. In J. L. Outtz (Ed.), *Adverse impact: Implications for organizational staffing and high stakes selection* (pp. 95–134). New York, NY: Routledge.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60,* 549–576. http://dx.doi.org/10.1146/annurev.psych.58.110405.085530

Greenman, E., Bodovski, K., & Reed, K. (2011). Neighborhood characteristics, parental practices and children's math achievement in elementary school. *Social Science Research, 40,* 1434–1444. http://dx.doi.org/10.1016/j.ssresearch.2011.04.007

Griggs v. Duke Power Co., 401 U.S. 424 (1971).

Heiland, F. (2009). Does the birth order affect the cognitive development of a child? *Applied Economics, 41,* 1799–1818. http://dx.doi.org/10.1080/00036840601083220

Higginbotham, A. L., Jr. (1980). *In the matter of color: Race and the American legal process. The colonial period* (Vol. 608). New York, NY: Oxford University Press.

Higginbotham, A. L., Jr. (1996). *Shades of freedom: Racial politics and presumptions of the American legal process* (Vol. 2). New York, NY: Oxford University Press.

Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection, and amelioration of adverse impact in personnel selection procedures: Issues, evidence, and lessons learned. *International Journal of Selection and Assessment, 9,* 152–194. http://dx.doi.org/10.1111/1468-2389.00171

Huang, Z. J., Lewin, A., Mitchell, S. J., & Zhang, J. (2012). Variations in the relationship between maternal depression, maternal sensitivity, and child attachment by race/ethnicity and nativity: Findings from a nationally representative cohort study. *Maternal and Child Health Journal, 16,* 40–50. http://dx.doi.org/10.1007/s10995-010-0716-2

Hunter, J. E., & Hunter, R. (1984). Validity and utility of alternative predictors. *Psychological Bulletin, 96,* 72–98. http://dx.doi.org/10.1037/0033-2909.96.1.72

Jaffe, L. E. (2009). *Development, interpretation, and application of the W score and the relative proficiency index* (Woodcock-Johnson III Assessment Service Bulletin No. 11). Rolling Meadows, IL: Riverside.

James, L. R., Mulaik, S. A., & Brett, J. M. (2006). A tale of two methods. *Organizational Research Methods, 9,* 233–244. http://dx.doi.org/10.1177/1094428105285144

Katznelson, I. (2005). *When affirmative action was White: An untold history of racial inequality in twentieth-century America*. New York, NY: Norton.

Kehoe, J. F. (2002). General mental ability and selection in private sector organizations: A commentary. *Human Performance, 15,* 97–106. http://dx.doi.org/10.1080/08959285.2002.9668085

Kelley, M. L., Power, T. G., & Wimbush, D. D. (1992). Determinants of disciplinary practices in low-income black mothers. *Child Development, 63,* 573–582. http://dx.doi.org/10.2307/1131347

Klebanov, P. K., Brooks-Gunn, J., & Duncan, G. J. (1994). Does neighborhood and family poverty affect mothers' parenting, mental health, and social support? *Journal of Marriage and the Family, 56,* 441–455. http://dx.doi.org/10.2307/353111

Klebanov, P. K., Brooks-Gunn, J., McCarton, C., & McCormick, M. C. (1998). The contribution of neighborhood and family income to developmental test scores over the first three years of life. *Child Development, 69,* 1420–1436. http://dx.doi.org/10.2307/1132275

Krieger, N., Rowley, D. L., Herman, A. A., Avery, B., & Phillips, M. T. (1993). Racism, sexism, and social class: Implications for studies of health, disease, and well-being. *American Journal of Preventive Medicine, 9*(Suppl. 6), 82–122.

Lemelin, J. P., Tarabulsy, G. M., & Provost, M. A. (2006). Predicting preschool cognitive development from infant temperament, maternal sensitivity, and psychosocial risk. *Merrill-Palmer Quarterly, 52,* 779–804. http://dx.doi.org/10.1353/mpq.2006.0038

Lhila, A., & Long, S. (2012). What is driving the Black–White difference in low birthweight in the US? *Health Economics, 21,* 301–315. http://dx.doi.org/10.1002/hec.1715

Linver, M. R., Brooks-Gunn, J., & Kohen, D. E. (2002). Family processes as pathways from income to young children's development. *Developmental Psychology, 38,* 719–734. http://dx.doi.org/10.1037/0012-1649.38.5.719

Lipsey, M. W., & Wilson, D. B. (2000). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Magnuson, K. (2007). Maternal education and children's academic achievement during middle childhood. *Developmental Psychology, 43,* 1497–1512. http://dx.doi.org/10.1037/0012-1649.43.6.1497

Mandara, J., Varner, F., Greene, N., & Richman, S. (2009). Intergenerational family predictors of the Black–White achievement gap. *Journal of Educational Psychology, 101,* 867–878. http://dx.doi.org/10.1037/a0016644

McKay, P. F., & McDaniel, M. A. (2006). A reexamination of Black–White mean differences in work performance: More data, more moderators. *Journal of Applied Psychology, 91,* 538–554. http://dx.doi.org/10.1037/0021-9010.91.3.538

McLoyd, V. C. (1990). The impact of economic hardship on black families and children: Psychological distress, parenting, and socioemotional development. *Child Development, 61,* 311–346. http://dx.doi.org/10.2307/1131096

Mills-Koonce, W. R., Appleyard, K., Barnett, M., Deng, M., Putallaz, M., & Cox, M. (2011). Adult attachment style and stress as risk factors for early maternal sensitivity and negativity. *Infant Mental Health Journal, 32,* 277–285. http://dx.doi.org/10.1002/imhj.20296

Mistry, R. S., Benner, A. D., Biesanz, J. C., Clark, S. L., & Howes, C. (2010). Family and social risk, and parental investments during the early childhood years as predictors of low income children's school readiness outcomes. *Early Childhood Research Quarterly, 25,* 432–449. http://dx.doi.org/10.1016/j.ecresq.2010.01.002

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.

Myrdal, G. (1944). *An American dilemma. Vol. 2*: *The Negro problem and modern democracy*. Piscataway, NJ: Transaction.

National Institute of Child Health and Human Development. (2005). *Child care and child development*. New York, NY: Guilford Press.

National Institute of Child Health and Human Development. (2012). Study Overview. Retrieved from http://www.nichd.nih.gov/research/supported/seccyd/pages/overview.aspx

Newman, D. A. (2003). Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods, 6,* 328–362. http://dx.doi.org/10.1177/1094428103254673

Newman, D. A. (2009). Missing data techniques and low response rates:

The role of systematic nonresponse parameters. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends* (pp. 7–36). New York, NY: Routledge.

Newman, D. A. (2014). Missing data: Five practical guidelines. *Organizational Research Methods, 17,* 372–411. http://dx.doi.org/10.1177/1094428114548590

Newman, D. A., & Cottrell, J. M. (2015). Missing data bias: Exactly how bad is pairwise deletion? In C. E. Lance & R. J. Vandenberg (Eds.), *More statistical and methodological myths and urban legends* (pp. 133–161). New York, NY: Routledge.

Newman, D. A., Hanges, P. J., & Outtz, J. L. (2007). Racial groups and test fairness, considering history and construct validity. *American Psychologist, 62,* 1082–1083. http://dx.doi.org/10.1037/0003-066X.62.9.1082

Newman, D. A., Jacobs, R. R., & Bartram, D. (2007). Choosing the best method for local validity estimation: Relative accuracy of meta-analysis versus a local study versus Bayes-analysis. *Journal of Applied Psychology, 92,* 1394–1413. http://dx.doi.org/10.1037/0021-9010.92.5.1394

Omi, M., & Winant, H. (1994). *Racial formation in the United States: From the 1960s to the 1990s* (2nd ed.). New York, NY: Routledge.

Outtz, J. L. (Ed.). (2010). *Adverse impact, implications for organizational staffing and high stakes selection*. New York, NY: Routledge.

Outtz, J. L., & Newman, D. A. (2010). A theory of adverse impact. In J. L. Outtz (Ed.), *Adverse impact: Implications for organizational staffing and high stakes selection* (pp. 95–134). New York, NY: Routledge.

Page, M., Wilhelm, M. S., Gamble, W. C., & Card, N. A. (2010). A comparison of maternal sensitivity and verbal stimulation as unique predictors of infant social-emotional and cognitive development. *Infant Behavior & Development, 33,* 101–110. http://dx.doi.org/10.1016/j.infbeh.2009.12.001

Peterson, G., & Peters, D. (1985). The socialization values of low-income Appalachian white and rural black mothers: A comparative study. *Journal of Comparative Family Studies, 16,* 75–91.

Phinney, J. S. (1996). When we talk about American ethnic groups, what do we mean? *American Psychologist, 51,* 918–927. http://dx.doi.org/10.1037/0003-066X.51.9.918

Ployhart, R. E., & Holtz, B. C. (2008). The diversity-validity dilemma: Strategies for reducing racioethnic and sex subgroup differences and adverse impact in selection. *Personnel Psychology, 61,* 153–172. http://dx.doi.org/10.1111/j.1744-6570.2008.00109.x

Popenoe, P. (1922). Intelligence and race: A review of some of the results of the army intelligence tests—II. The negro. *The Journal of Heredity, 13,* 295–298.

Pungello, E. P., Iruka, I. U., Dotterer, A. M., Mills-Koonce, R., & Reznick, J. S. (2009). The effects of socioeconomic status, race, and parenting on language development in early childhood. *Developmental Psychology, 45,* 544–557. http://dx.doi.org/10.1037/a0013917

Raby, K. L., Roisman, G. I., Fraley, R. C., & Simpson, J. A. (in press).The enduring predictive significance of early maternal sensitivity: Social and academic competence through age 32 years. *Child Development*. Advance online publication. http://dx.doi.org/10.1111/cdev.12325

Rauh, V. A., Achenbach, T. M., Nurcombe, B., Howell, C. T., & Teti, D. M. (1988). Minimizing adverse effects of low birthweight: Four-year results of an early intervention program. *Child Development, 59,* 544–553. http://dx.doi.org/10.2307/1130556

Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology, 54,* 297–330. http://dx.doi.org/10.1111/j.1744-6570.2001.tb00094.x

Roth, P. L., Huffcutt, A. I., & Bobko, P. (2003). Ethnic group differences in measures of job performance: A new meta-analysis. *Journal of Applied Psychology, 88,* 694–706.

Sackett, P. R., & Ellingson, J. E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology, 50,* 707–721. http://dx.doi.org/10.1111/j.1744-6570.1997.tb00711.x

Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary academic performance? *Psychological Bulletin, 135,* 1–22. http://dx.doi.org/10.1037/a0013978

Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative-action world. *American Psychologist, 56,* 302–318. http://dx.doi.org/10.1037/0003-066X.56.4.302

Sackett, P. R., & Shen, W. (2010). Subgroup differences on cognitively loaded tests in contexts other than personnel selection. In J. L. Outtz (Ed.), *Adverse impact: Implications for organizational staffing and high stakes selection* (pp. 323–346). New York, NY: Routledge.

Santrock, J. W. (1972). Relation of type and onset of father absence to cognitive development. *Child Development, 43,* 455–469. http://dx.doi.org/10.2307/1127548

Schady, N. (2011). Parents' education, mothers' vocabulary, and cognitive development in early childhood: Longitudinal evidence from Ecuador. *American Journal of Public Health, 101,* 2299–2307. http://dx.doi.org/10.2105/AJPH.2011.300253

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7,* 147–177.

Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin, 64,* 92–107. http://dx.doi.org/10.1037/h0022371

Scherbaum, C. A., Goldstein, H. W., Yusko, K. P., Ryan, R., & Hanges, P. J. (2012). Intelligence 2.0: Reestablishing a research program on *g* in I-O psychology. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 5,* 128–148. http://dx.doi.org/10.1111/j.1754-9434.2012.01419.x

Schmidt, F. L., & Hunter, J. E. (1981). Employment testing: Old theories and new research findings. *American Psychologist, 36,* 1128–1137. http://dx.doi.org/10.1037/0003-066X.36.10.1128

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124,* 262–274. http://dx.doi.org/10.1037/0033-2909.124.2.262

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology, 86,* 162–173. http://dx.doi.org/10.1037/0022-3514.86.1.162

Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology, 71,* 432–439. http://dx.doi.org/10.1037/0021-9010.71.3.432

Schmitt, N., & Quinn, A. (2010). Reduction in measured subgroup mean differences: What is possible? In J. L. Outtz (Ed.), *Adverse impact: Implications for organizational staffing and high stakes selection* (pp. 95–134). New York, NY: Routledge.

Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings, D. (1997). Adverse impact and predictive efficiency of various predictor combinations. *Journal of Applied Psychology, 82,* 719–730. http://dx.doi.org/10.1037/0021-9010.82.5.719

Shinn, M. (1978). Father absence and children's cognitive development. *Psychological Bulletin, 85,* 295–324. http://dx.doi.org/10.1037/0033-2909.85.2.295

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195152968.001.0001

Society for Industrial and Organizational Psychology. (2013, November 30). Adverse Impact Reduction Research Initiative and Action (AIRRIA) Research Grant. Retrieved from http://www.siop.org/SIOPAwards/adverse.aspx

Stams, G. J., Juffer, F., & van IJzendoorn, M. H. (2002). Maternal sensitivity, infant attachment, and temperament in early childhood predict adjustment in middle childhood: The case of adopted children and their biologically unrelated parents. *Developmental Psychology, 38,* 806–821. http://dx.doi.org/10.1037/0012-1649.38.5.806

Svanum, S., Bringle, R. G., & McLaughlin, J. E. (1982). Father absence and cognitive performance in a large sample of six- to eleven-year-old children. *Child Development, 53,* 136–143. http://dx.doi.org/10.2307/1129645

Taylor, H. G., Klein, N., Minich, N. M., & Hack, M. (2000). Middle-school-age outcomes in children with very low birthweight. *Child Development, 71,* 1495–1511. http://dx.doi.org/10.1111/1467-8624.00242

Terrill, W., & Reisig, M. D. (2003). Neighborhood context and police use of force. *Journal of Research in Crime and Delinquency, 40,* 291–321. http://dx.doi.org/10.1177/0022427803253800

Thompson, R. J., Jr., Catlett, A. T., Oehler, J. M., Gustafson, K. E., Goldstein, R. F., & Prochaska, J. J. (1998). Home environment and developmental outcome of African American and White infants with very low birthweight. *Children's Health Care, 27,* 1–14. http://dx.doi.org/10.1207/s15326888chc2701_1

Torche, F., & Echevarría, G. (2011). The effect of birthweight on childhood cognitive development in a middle-income country. *International Journal of Epidemiology, 40,* 1008–1018. http://dx.doi.org/10.1093/ije/dyr030

Travis, R., & Kohli, V. (1995). The birth order factor: Ordinal position, social strata, and educational achievement. *The Journal of Social Psychology, 135,* 499–507. http://dx.doi.org/10.1080/00224545.1995.9712218

United States Census Bureau. (2009a). Table 695: Money income of families—Number and distribution by race and Hispanic origin: 2009. Retrieved from http://www.census.gov/compendia/statab/2012/tables/12s0695.pdf

United States Census Bureau. (2009b). Table 696: Money income of families—Percent distribution by income level in constant (2009) dollars: 1980 to 2009. Retrieved from http://www.census.gov/compendia/statab/2012/tables/12s0695.pdf

United States Census Bureau. (2012a). Table 230: Educational attainment, by race, Hispanic origin, and sex Retrieved from http://www.census.gov/compendia/statab/2012/tables/12s0230.pdf

United States Census Bureau. (2012b). Poverty thresholds by size of family and number of children. Retrieved from https://www.census.gov/hhes/www/poverty/data/threshld/thresh12.xlsx

Waldman, D. A., & Avolio, B. J. (1991). Race effects in performance evaluations: Controlling for ability, education, and experience. *Journal of Applied Psychology, 76,* 897–901. http://dx.doi.org/10.1037/0021-9010.76.6.897

Watson, J. E., Kirby, R. S., Kelleher, K. J., & Bradley, R. H. (1996). Effects of poverty on home environment: An analysis of three-year outcome data for low birth weight premature infants. *Journal of Pediatric Psychology, 21,* 419–431. http://dx.doi.org/10.1093/jpepsy/21.3.419

Wee, S., Newman, D. A., & Joseph, D. L. (2014). More than g: Selection quality and adverse impact implications of considering second-stratum cognitive abilities. *Journal of Applied Psychology, 99,* 547–563. http://dx.doi.org/10.1037/a0035183

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin, 116,* 363–381. http://dx.doi.org/10.1037/0033-2909.116.2.363

Williams, T., & Kornblum, W. (1985). *Growing up poor*. Lexington, MA: Lexington Books.

Wilson, H. (1974). Parenting in poverty. *British Journal of Social Work, 4,* 241–254.

Winant, H. (2000). Race and race theory. *Annual Review of Sociology, 26,* 169–185. http://dx.doi.org/10.1146/annurev.soc.26.1.169

Woodcock, R. W. (1990). Theoretical foundations of the WJ-R measures of cognitive ability. *Journal of Psychoeducational Assessment, 8,* 231–258. http://dx.doi.org/10.1177/073428299000800303

Woodcock, R. W., & Dahl, M. N. (1971). *A common scale for the measurement of person ability and test item difficulty* (AGS Paper No. 10). Circle Pines, MN: American Guidance Service.

Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock–Johnson Psycho-Educational Battery—Revised*. Allen, TX: Developmental Learning Materials.

Yeung, W. J., & Conley, D. (2008). Black–White achievement gap and family wealth. *Child Development, 79,* 303–324. http://dx.doi.org/10.1111/j.1467-8624.2007.01127.x

Yeung, W. J., Linver, M. R., & Brooks-Gunn, J. (2002). How money matters for young children's development: Parental investment and family processes. *Child Development, 73,* 1861–1879. http://dx.doi.org/10.1111/1467-8624.t01-1-00511

Yeung, W. J., & Pfeiffer, K. M. (2009). The Black–White test score gap and early home environment. *Social Science Research, 38,* 412–437. http://dx.doi.org/10.1016/j.ssresearch.2008.11.004

Yusko, K. P., & Goldstein, H. W. (2008). *Siena Reasoning Test*. Princeton, NJ: Siena Consulting.

Zajonc, R. B., & Bargh, J. A. (1980). Birth order, family size, and decline of SAT scores. *American Psychologist, 35,* 662–668. http://dx.doi.org/10.1037/0003-066X.35.7.662

Zajonc, R. B., & Markus, G. B. (1975). Birth order and intellectual development. *Psychological Review, 82,* 74–88. http://dx.doi.org/10.1037/h0076229

Zajonc, R. B., & Mullally, P. R. (1997). Birth order: Reconciling conflicting effects. *American Psychologist, 52,* 685–699. http://dx.doi.org/10.1037/0003-066X.52.7.685

Zajonc, R. B., & Sulloway, F. J. (2007). The confluence model: Birth order as a within-family or between-family dynamic? *Personality and Social Psychology Bulletin, 33,* 1187–1194. http://dx.doi.org/10.1177/0146167207303017

Zedeck, S. (2010). Adverse impact: History and evolution. In J. L. Outtz (Ed.), *Adverse impact: Implications for organizational staffing and high stakes selection* (pp. 3–27). New York, NY: Routledge.

(*Appendices follows*)

Appendix A

Past Studies Attempting to Explain the Black-White Gap in Cognitive Test Scores

Authors	Sample type	Sample size	Ability measures	Explanatory variables	Race–*g* relationship	Statistically significant explanatory variables
Brooks-Gunn, Klebanov, Smith, Duncan, & Lee (2003)	Study 1: Two-wave, ages 3 and 5, nationally representative from Infant Health and Development Program (IHDP) Study 2: Cross-sectional, ages 3–4 and 5–6, restricted to low-birth-weight children from National Longitudinal Study of Youth-Child Supplement (NLSY-CS)	N = 627 (312 Black, 315 White) N = 2,220 for 3–4 years old, N = 1,354 for 5–6 years old	IHDP: (PPVT-R) ages 3 and 5, Stanford-Binet Intelligence Test age 3, (WPPSI) age 5 NLSY-CS: PPVT-R for children, AFQT for mothers Measurement equivalence over time not assessed Slope differences not assessed	Birth weight Gender Family income Female head of household Maternal education Maternal verbal ability Maternal age HOME Learning HOME Warmth	IHDP: Standardized regression coefficient drops from an average of $-.49^{*}$ to $-.19^{*}$ when all covariates are included NLSY-CS: Standardized regression coefficient drops from an average of $-.49^{*}$ to $-.30^{*}$ when all covariates are included	IDHP PPVT-R Age 5: Income Maternal education HOME Learning HOME Warmth NLSY-CS PPVT-R Age 5: Gender Birth weight Female head of household Maternal education Maternal verbal ability HOME Learning HOME Warmth Covariates not reported for Stanford-Binet and Wechsler
Fryer & Levitt (2004)	Longitudinal (4 time points, fall and spring of kindergarten, spring of first grade, subsample for fall of first grade), ECLS nationally representative both public and private, full-time and part-time schools and kindergartens	N = 13,290 for math N = 12,601 for reading	Math and reading tests developed exclusively for ECLS Measurement equivalence over time not assessed Slope differences not assessed Separate regression model at each time point	Models 4 & 9, p. 451: (composite of parental education, occupational status, & household income) Number of children's books Number of children's books squared Birth weight Mother over 30 at first birth Teenage mother at first birth Gender Child age at kindergarten Participation in nutrition program (Women, Infants, and Children [WIC])	Math: Unstandardized regression coefficient for Black–White gap at fall of kindergarten reduced from $-.638^{*}$ to $-.094^{*}$ with covariates included Reading: Unstandardized regression coefficient for Black–White gap reduced from $-.401^{*}$ to $+.117^{*}$ with covariates included Spring first-grade gap (math): $b = -.250^{*}$; (reading): $b = -.071^{*}$	SES Number of children's books Number of children's books squared Gender (reading only) Child age at kindergarten Birth weight Mother over 30 at first birth Teenage mother at first birth Participant in nutrition program (WIC) Every covariate significant at fall of kindergarten was significant at spring of first grade

(*Appendices continue*)

Authors	Sample type	Sample size	Ability measures	Explanatory variables	Race–*g* relationship	Statistically significant explanatory variables
Fryer & Levitt (2006)	Longitudinal (4 time points, fall of kindergarten, spring of kindergarten, spring of first grade, spring of third grade) using data from the Early Childhood Longitudinal Study (ECLS)	Total *N*: 11,201 for math, 10,540 for reading	Math and reading tests developed exclusively for ECLS based on existing instruments Measurement equivalence over time not assessed Slope differences not assessed Separate regression model at each time point	SES (same as Fryer & Levitt, 2004) Number of children's books Number of children's books squared Birth weight Mother over 30 at first birth Teenage mother at first birth Gender Child age at kindergarten Participation in nutrition program (WIC)	Math: unstandardized regression coefficient (average over 4 time points) reduced from .76* to .24* with covariates included Reading: unstandardized regression coefficient (average over 4 time points) drops from .53* to .06* with covariates included	Child age at kindergarten Birth weight Gender (all reading, first- and third-grade math) Number of children's books Number of children's books squared Mother over 30 at first birth SES Participation in nutrition program (WIC) Teenage mother at first birth
Mandara, Varner, Greene, & Richman (2009)	Two-wave, ages 10–11 and 13–14, with parents' data included, data from 1978 NLSY	$N = 4{,}406$ children, 2,284 mothers	AFQT for parents Three PIAT subtests for children (reading recognition, reading comprehension, mathematical reasoning) Measurement equivalence over time not assessed Slope differences not assessed	Grandparent SES (occupational prestige, education, library resources) Mother's achievement test scores Family SES (occupational prestige, poverty status, wealth) Child decision making Parental monitoring of children Child house chores Arguing about rules School-oriented home Maternal warmth	Black–White $d = .81$ (arithmetic reasoning) $d = .62$ (word recognition) $d = .75$ (reading comprehension) Drops to overall standardized regression coefficient $\beta = -.07^*$ (favoring Black)	Grandparent SES Mother's achievement test scores Family SES Child decision making Child house chores Arguing about rules

(*Appendices continue*)

Authors	Sample type	Sample size	Ability measures	Explanatory variables	Race–*g* relationship	Statistically significant explanatory variables
Yeung & Pfeiffer (2009)	Two-wave, 3 groups. First cohort is kindergarten in 1997, Grades 4–6 in 2003, second cohort is Grades 1–3 in 1997, Grades 7–9 in 2003, and third group is Grades 4–7 in 1997, Grades 10–12 in 2003; All are from PSID	N = 1,794 (856 Black and 938 White) across the three cohorts	WJ-R Applied Problems subtest and Letter-Word Identification subtest (tests are age standardized) Passage comprehension test for mothers Subtests analyzed separately Measurement equivalence over time not assessed Slope differences not assessed	Gender Paternal grandparent education Maternal grandparent education Mother received federal aid when child was born Teenage mother Low birth weight Birth order Parental education Parental occupational prestige Income from birth to age 5 Average family wealth Number of children Family structure Urbanicity Parental expectations Cognitive stimulation Emotional support at home Weekly TV time Mother's verbal test score	Cohort 1, preschool, 1997: Math: Gap drops from unstandardized regression coefficient of –.78* to –.24 Reading: gap drops from –.43* to .02. Cohort 1, Grades 4–6 2003: Math: Gap drops from –.98* to –.43* Reading: gap drops from –.67* to .02 Cohort 2, Grades 1–3, 1997: Math: Gap drops from –.67* to –.10 Reading: Gap drops from .84* to –.10 Cohort 2, Grades 7–9, 2003: Math: Gap drops from –1.0* to –.47* Reading: gap drops from –.94* to −.41* Cohort 3: Grades 4–7, 1997: Math: Gap drops from –.77* to –.47* Reading: gap drops from –.77* to –.22 Cohort 3: Grades 10–12, 2003: Math: Gap drops from –.78* to –.58* Reading: Gap drops from –.74* to –.40*	Teenage mother Birth order Low birth weight Occupational prestige Income from birth to age 5 Parental expectations Mother's verbal test score Urbanicity Weekly TV time Parental education Net wealth Number of children Gender

(*Appendices continue*)

Authors	Sample type	Sample size	Ability measures	Explanatory variables	Race–*g* relationship	Statistically significant explanatory variables
Burchinal et al. (2011)	NICHD SECCYD Longitudinal (4 time points: 54 months, first grade, third grade, fifth grade) Low-income sample only (2.25× poverty line and below); Dropped over 400 participants who were above poverty threshold	$N = 314$	Woodcock-Johnson Revised (WJ-R) Applied problems at 54 months and first grade Letter-Word ID at 54 months and first grade Broad Reading at third and fifth grades Broad Math at third and fifth grades Reading and math analyzed separately Reading operationalized differently at T1 and T2 versus T3 and T4 Math operationalized differently at T1 and T2 versus T3 and T4 Tested intercept and slope differences over time	Gender Maternal education Whether child was firstborn Maternal childrearing attitudes One- or two-parent household Income-to-needs ratio Parenting composite (age standardized composite average of HOME ratings and maternal sensitivity ratings) Neighborhood disadvantage (Census block indices of household income, employment status, marital status) Site (of hospital) School risk (proportion of student body receiving free or reduced price lunch and non-White proportion of student body) Classroom quality (observer ratings) Regression models include interaction terms of every covariate with age Only statistically significant regression coefficients are reported	Reading, intercept: Unstandardized regression coefficient drops from –12.53* to –3.80 Reading, slope: Unstandardized regression coefficient increases from –.40 to –.66 Math, intercept: Unstandardized regression coefficient drops from –9.69* to 1.08 Math, slope: Unstandardized regression coefficient drops from .97* to .16 Black–White intercept differences are significant for math and reading in favor of White Black–White slope differences are significant only for math in favor of Black	Intercept: Parenting quality composite Whether child was firstborn (reading only) Neighborhood disadvantage (math only) Child–teacher ratio (math, Black only) Slope: Two-parent household (reading only) Classroom quality (math, Black only) Gender (math only)

(*Appendices continue*)

Authors	Sample type	Sample size	Ability measures	Explanatory variables	Race–*g* relationship	Statistically significant explanatory variables
Current article	NICHD (SECCYD) Longitudinal (5 time points: 54 months, first grade, third grade, fifth grade, 15 years), full income range, families recruited from U.S. hospitals	$N = 791$	WJ-R Applied problems (math) at all time points Letter-Word ID (reading) at 54 months, first grade, third grade, and fifth grade Passage comprehension (reading) at 15 years Picture-Vocabulary at all time points Estimated measurement model for *g* Assessed measurement equivalence over time Tested *g* intercept and slope differences over time	Birth order Maternal verbal ability/knowledge Learning materials Maternal sensitivity Maternal warmth and acceptance Physical environment Birth weight Income-to-needs ratio Maternal education	Overall *g* gap drops from standardized beta of –.42* to –.07 with covariates included	Birth order Maternal verbal ability/knowledge Learning materials Maternal sensitivity Physical environment

Note. IHDP = Infant Health and Development Program; NLSY-CS = National Longitudinal Study of Youth - Child Supplement; ECLS = Early Childhood Longitudinal Study; PPVT-R = Peabody Picture Vocabulary Test - Revised; WPPSI = Wechsler Preschool and Primary Scale of Intelligence; AFQT = Armed Forces Qualifying Tests; SES = Socioeconomic status; WIC = Women, Infants, and Children; HOME = Home Observation for Measurement of the Environment; PIAT = Peabody Individual Achievement Test; PSID = Panel Study of Income Dynamics; WJ-R = Woodcock-Johnson Revised; NICHD = National Institute of Child Health and Human Development; SECCYD = Study of Early Child Care and Youth Development; WJ-R = Woodcock-Johnson Revised; T1 = Time 1; T2 = Time 2; T3 = Time 3; T4 = Time 4.

(*Appendices continue*)

Appendix B

Latent Growth Model Specifications and Results

Growth Model Specifications

A sequence of four a priori models was specified, and estimated in Mplus Version 7 (Muthén & Muthén, 2012). The first model was a latent growth model (LGM) for changes in cognitive ability/knowledge (*g*) over time (Model 1a: Cognitive Test LGM). At each time point, *g* had three indicators: math, vocabulary, and reading tests (to set the scale, loadings of the math test onto *g* at each time point in time were fixed to 1.0). A *g intercept factor* was created, onto which the first-order *g* factors from each time point had a fixed loading of 1.0. To create a *g slope factor*, we fixed the loadings of the *g* factors to increase linearly with time (loadings were fixed to −0.1 [T1: 54 month *g*], 0.1 [T2: first grade *g*], 0.3 [T3: third grade *g*], 0.5 [T4: fifth grade *g*], and 1.0 [T5: 15 year/10th grade *g*]). These values were chosen to approximate the growth of the child over the available time points (for example, −0.1 is before formal schooling begins [before kindergarten], .1 is first grade, .3 is third grade, .5 is fifth grade, and so forth). Additionally, uniquenesses for each of the indicators of *g* (e.g., for math test scores) were allowed to correlate over time (that is, we freed all autoregressive error covariances, within each indicator [Singer & Willett, 2003]; average autoregressive correlation = .16, range = .09 to .39). The second model was identical to the first model, except (as recommended by Chan, 1998) when implementing the growth model, we first sought to establish measurement equivalence over time. That is, math, vocabulary, and reading loadings onto *g* were constrained to be invariant over time (Model 1b: Cognitive Test LGM, with Measurement Equivalence).

The third model is an LGM for *g*, which is identical to the second model (Model 1b), but with the addition of parameters to estimate racial differences in both the *g* intercept and *g* slope (Model 2: Race and Cognitive Test LGM). The fourth model includes the explanatory variables that we hypothesized would account for the association between race and cognitive test scores (Model 3: Race, Explanatory Variables, and Cognitive Test LGM). Factors were specified for birth order, maternal cognitive test scores, learning materials, maternal acceptance, physical environment, and birth weight, with each reflected by a single manifest indicator with a fixed loading of 1.0 and uniqueness of zero. For maternal sensitivity and income (which were each measured at five points in time), we modeled the time intercept by fixing the loadings of maternal sensitivity and income from each time point to 1.0 (Willett & Sayer, 1994). In this way, the time intercept factors for income and maternal sensitivity are methodologically similar to the *g* intercept (which they are intended to explain). These time intercept factors capture the portion of variance in income, maternal sensitivity, and *g* that does not change over time. Although some covariates (maternal sensitivity and income) were measured at five points in time leading up to 54 months of age and were specified as time intercepts, other covariates (learning materials, acceptance, and physical environment) used single-time-point measures taken at 54 months of age. We consider the 54-month timing of these covariates to be appropriate for our current purposes, because our objective is to explain Black–White differences in the *g* intercept, which already existed at 54 months of age (i.e., just before participants entered school).

Measurement Model Results for Cognitive Test Scores and Race

Model 1a is a measurement model for *g* over time. Overall fit for Model 1a is adequate, $\chi^2(57) = 225.69$, Comparative Fit Index (CFI) = .98, Non-Normed Fit Index (NNFI) = .97, Root Mean Squared Error of Approximation (RMSEA) = .061, and all standardized factor loadings for Model 1a are greater than .60. Model 1b is identical to Model 1a, but with measurement equivalence specified over time (Chan, 1998), such that the loadings of each of the three indicators of *g* (math, vocabulary, and reading tests) are fixed to be equal over time. Fit of this constrained Model 1b is also deemed adequate, $\chi^2(65) = 313.29$, CFI = .98, NNFI = .96, RMSEA = .069. Because the change in CFI between Model 1a and Model 1b was less than .01 (Cheung & Rensvold, 2002), we interpret these results to suggest adequate measurement equivalence over time, enabling subsequent longitudinal modeling.

Table B1
Structural Equation Modeling Results Involving Covariates

Predictor variables	DV: *g* intercept		Each covariate alone: Indirect effect size (Percent of race gap explained)	Full model: Indirect effect size (Percent of race gap explained)
	Model 2	Model 3		
Race (W = 0; B = 1)	−.42*	−.07		
Maternal sensitivity		.21*	−.267* (61.9%)	−.108* (25.2%)
Maternal acceptance		.04	−.073* (17.2%)	−.011 (2.7%)
Physical environment		.08*	−.089* (20.7%)	−.023* (5.4%)
Learning materials		.08*	−.126* (29.3%)	−.029* (6.8%)
Birth order		−.17*	−.026* (6.1%)	−.022* (5.1%)
Birth weight		.03	−.013 (3.1%)	−.006 (1.5%)
Maternal verbal ability/knowledge		.33*	−.238* (54.6%)	−.142* (33.5%)
Income		.01	−.098* (22.8%)	−.002 (0.5%)
Maternal education		.06	−.103* (23.8%)	−.014 (3.2%)

Note. $N = 791$. Coefficients are standardized. DV = dependent variable; W = White; B = Black.
* $p < .05$.

(*Appendices continue*)

For Model 2, we added race to the measurement model (i.e., we added race to Model 1b) and allowed race to relate to the *g* intercept and *g* slope (a figure displaying all Model 2 parameter estimates is available from the first author). Adding race to the model results in adequate model fit, $\chi^2(78) = 390.94$, CFI = .97, NNFI = .96, RMSEA = .071. In Model 2, the race–*g intercept* path coefficient was large and statistically significant (standardized $\gamma = -0.42$, $p < .05$, $d = -1.35$). In contrast, the path coefficient between race and the *g slope* was not statistically significant (standardized $\gamma = -0.05$, $p > .05$, *ns*, $d = -.11$).

Using Explanatory Variables to Account for the Cognitive Test Score Gap

Model 3 expands Model 2 by adding the explanatory variables. In Model 3, we only specify relations between the explanatory variables and the *g intercept*. The *g slope* was not related to race, and therefore no explanatory variables were needed (because there was no race gap in the *g* slope).[3] Thus, we henceforth refer to the *g* intercept as simply "*g*."

Model 3 displays adequate fit to the data, $\chi^2(430) = 1185.19$, CFI = .96, NNFI = .95, RMSEA = .047. Table B1 shows the race results from Models 2 and 3 (race was coded as White = 0, Black = 1; therefore, a positive relation between a variable and race means the variable has a higher mean for Black participants; see Table 2 in the main text for the *d* values). In Model 3, the effect of race on *g* was no longer statistically significant ($\gamma = -0.07$, $p > .05$, *ns*) after all explanatory variables were added.[4]

We also tested the indirect effect of each explanatory variable as an explanation for the relation between race and cognitive scores. That is, we attempt to estimate the extent to which each explanatory variable accounts for (or explains) the race/test-score gap. In Table B1, we first present the indirect effects for each explanatory variable tested independently, and then we present the indirect effects for each explanatory variable from the full model (Model 3), in which all explanatory variables were tested simultaneously. The full model allows us to partition the race/test-score total effect (total race gap) into portions of the gap that were accounted for by each explanatory variable. When each explanatory variable is considered alone, the variables that best account for the indirect effect from race to *g* are (in order by magnitude; see third column of Table B1): maternal sensitivity (61.9% of the race gap in *g* explained), maternal verbal ability/knowledge (54.6% of the race gap explained), learning materials (29.3% of the race gap explained), maternal education (23.8% of the race gap explained), family income-to-needs (22.8% of the race gap explained), physical environment (20.7% of the race gap explained), maternal warmth and acceptance (17.2% of the race gap explained), and birth order (6.1% of the race gap explained). However, in the full model (Model 3), the explanatory variables that do the best job of *uniquely* accounting for the indirect effect from race to *g* are (in order by magnitude; see last column of Table B1): maternal verbal ability/knowledge (33.5%), maternal sensitivity (25.2%), learning materials (6.8%), physical environment (5.4%), and birth order (5.1%). Altogether, the set of explanatory variables in Model 3 accounts for 83.9% of the total race gap in cognitive test scores.

[3] At a reviewer's request, we tested several Time × Explanatory Variable interaction effects (e.g., Maternal Sensitivity × Age). Results of this analysis for each of the nine explanatory variables showed that none was moderated by age (i.e., the explanatory variables generalized across time).

[4] At a reviewer's request, we examined Race × Explanatory Variable interaction effects (e.g., Race × Birth Order) in separate models to predict both the *g* intercept and *g* slope. None of the 18 tests (i.e., nine Explanatory Variable × Race interactions predicting *g* intercept and slope) was significant.

Received December 23, 2013
Revision received January 11, 2015
Accepted January 24, 2015 ■