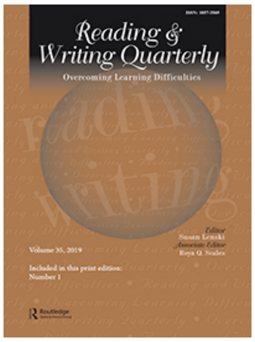


Routledge
Taylor & Francis Group

Reading & Writing Quarterly

Overcoming Learning Difficulties

ISSN: 1057-3569 (Print) 1521-0693 (Online) Journal homepage: https://www.tandfonline.com/loi/urwl20

Reading and Language Performance of Low-Income, African American Boys in Grades 1–5

Julie A. Washington, Lee Branum-Martin, Ryan Lee-James & Congying Sun

To cite this article: Julie A. Washington, Lee Branum-Martin, Ryan Lee-James & Congying Sun (2019) Reading and Language Performance of Low-Income, African American Boys in Grades 1–5, Reading & Writing Quarterly, 35:1, 42-64, DOI: 10.1080/10573569.2018.1535777

To link to this article: https://doi.org/10.1080/10573569.2018.1535777

Published online: 08 Jan 2019.

READING & WRITING QUARTERLY
2019, VOL. 35, NO. 1, 42–64
https://doi.org/10.1080/10573569.2018.1535777

Reading and Language Performance of Low-Income, African American Boys in Grades 1–5

Julie A. Washington[1], Lee Branum-Martin[1], Ryan Lee-James[2], and Congying Sun[1]

[1]Georgia State University, Atlanta, Georgia, USA; [2]Adelphi University, Garden City, New York, USA

ABSTRACT

This investigation examined the gender gap in language and reading skills in a sample of low-income African American boys compared to African American girls from the same neighborhoods and schools. Using a longitudinal, accelerated cohort design, we used individual growth curve models to evaluate the reading and language performance of 1st through 5th graders. We analyzed data for 7 outcomes: (a) language, (b) letter-word identification, (c) passage comprehension, (d) decoding, (e) reading fluency, (f) reading vocabulary, and (g) intelligence. Descriptive statistics revealed no statistically significant differences in performance on language or intelligence measures between boys and girls at any grade level. Conversely, all 5 reading skills measured (e.g., decoding, fluency, and reading vocabulary) showed significant differences in performance by gender favoring girls, but only in 4th and 5th grades. Growth models revealed no differences in the growth trajectories of boys and girls for language or intelligence. However, we observed gender differences in growth trajectories for 2 reading measures—passage comprehension and reading fluency—with girls demonstrating slightly faster growth compared to boys on a measure of passage comprehension and boys showing significant deceleration in reading fluency at 5th grade. We discuss the results relative to African American boys and expected patterns of achievement in language and reading during elementary school.

Individual differences in the performance of young children can be influenced by a host of sociodemographic and environmental factors, including socioeconomic status (SES), race, culture, gender, age, and quality of the home and school environments. These factors frequently interact to influence various child outcomes. The role of gender in the performance of boys compared to girls has been examined across several skill areas with mixed results. In the case of language and reading, it has been widely reported that girls develop a range of linguistic skills at an accelerated rate and at younger ages than do boys (Bauer, Goldfield, & Reznick, 2002; Eriksson et al., 2012; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Hyde & Linn, 1988; Maccoby & Jacklin, 1974) and that they also outpace their male peers in the development of early reading skills (Caro, McDonald, & Willms, 2009; Flannery, Liederman, Daly, & Schultz, 2000; Liederman, Kantrowitz, & Flannery, 2005; Logan & Johnston, 2010; McGeown, Goodwin, Henderson, & Wright, 2012). Despite seemingly strong and well-accepted evidence of the presence of gender differences, many empirical studies challenge these findings, showing either no gender differences for boys compared to girls in these important skill areas (Shaywitz, Shaywitz, Fletcher, & Escobar, 1990) or reporting strengths for boys over girls (Kidd & Lum, 2008; Logan & Johnston, 2010).

CONTACT Julie A. Washington jwashington@gsu.edu College of Education and Human Development, Georgia State University, P.O. Box 3979, Atlanta, GA 30302-3739, USA.

Gender differences in the development of reading and language skills have also been reported by race. In particular, African American boys are reportedly at highest risk for the development of poor outcomes in these and other achievement areas. Utilizing data from the National Assessment of Educational Progress (NAEP), Husband (2012b) and others (Davis, 2003; Jencks & Phillips, 2011) have discussed substantial race gaps in reading performance nationally, with African American boys performing significantly lower than Asian, White, Hispanic, and American Indian boys at fourth grade. In the early years the language and reading gaps between African American and White boys are reportedly evident as young as toddler age for language and preschool age for reading (Aratani, Wight, & Cooper, 2011). However, when researchers control for influential variables such as SES and child-level variables such as low birth weight and family support, the race gap for Black boys reportedly disappears by kindergarten (Aratani et al., 2011; Iruka, 2017). These findings are in contrast to several studies that report substantial gaps in achievement for African American boys throughout their schooling experience regardless of SES (Fantuzzo, LeBoeuf, Rouse, & Chen, 2012).

Though many studies highlight a gender gap in language and reading in early childhood, the findings are equivocal for school-age children. It is evident that something important happens between kindergarten and 12th grade when the performance of African American students in general and African American boys in particular reportedly diverges sharply from the performance of their peers. There are very few studies focused on this Race × Gender interaction in the achievement of school-age African American boys.

Many of these studies are cross-sectional and vary widely in age and grade ranges examined, sample size, and academic and cognitive skills assessed. In many of these studies, reading is represented as a combined achievement outcome, either as a composite score along with other associated academic skills (e.g., critical thinking, distinguishing real and imaginary text, and the ability to connect text with personal life experiences) without disaggregation of core reading skills (e.g., decoding, fluency, and comprehension; Matthews, Kizzie, Rowley, & Cortina, 2010) or as a criterion-based, pass/fail outcome (i.e., failing in one or more content areas, such as reading, math, or both; Connell, Spencer, & Aber, 1994), which makes it difficult to isolate the effect of gender on reading. The purpose of the current study was to characterize the development of language, reading, and cognition in African American boys in first through fifth grades utilizing a longitudinal, accelerated cohort design in an effort to contribute to understanding of the growth trajectories of these critical skills during elementary grades, a foundational time in schooling.

Gender differences in language

It is a long-held belief that boys develop language at a different rate than girls, with girls being superior to boys in overall verbal development (Anastasi, 1958; Hyde & Linn, 1988; Maccoby & D'Andrade, 1966; Maccoby & Jacklin, 1974). In particular, research in early childhood indicates that boys lag behind girls in their development of vocabulary, gestures, word combining, utterance length, and complexity (Bauer et al., 2002; Eriksson et al., 2012; Huttenlocher et al., 1991). In an investigation of gender-based language differences in infancy and toddlerhood across 10 non-English language communities, Eriksson et al. (2012) demonstrated that beginning as early as infant communication, including gestures, girls were ahead of boys developmentally, though effect sizes were quite small, accounting for <1% of the variance in performance.

Similarly, Bauer et al. (2002) and Huttenlocher et al. (1991) found significant differences in lexical development between boys and girls who were 2 years of age or younger. Though these gender differences are consistently evident in early childhood, by the time children reach school age, the language gap between boys and girls reportedly disappears (Maccoby & D'Andrade, 1966; Maccoby & Jacklin, 1974). It is important to note that because of the young ages of children in these investigations, parent-report measures are most often used to estimate the size of children's vocabulary and the extent of their expressive and receptive language skills. Though this is an appropriate methodology for use with very young children, self-report has widely acknowledged limitations in terms of accurate characterization of gender differences in early language abilities.

Studies of gender differences in school-age children have reported less consistent conclusions, and fewer recent articles are available. In an early meta-analysis of studies focused on gender differences in language ability, Maccoby and Jacklin (1974) concluded that the preponderance of the evidence suggested that the language development rate and skills of boys and girls were very similar from preschool through preadolescence. At approximately 10 or 11 years of age there is a divergence, such that girls outperform boys linguistically, and this advantage continues through high school. According to Maccoby and Jacklin (1974), the magnitude of the gender difference in performance can vary but amounts to approximately 0.25 *SD*.

Most important perhaps is that of 85 studies examined, 70% reported no gender differences at all in verbal abilities, 25% reported an advantage for females, and the remaining 5% reported an advantage for the males in their samples. These studies varied widely in terms of sample sizes, verbal abilities assessed, and the age ranges of participants. In stark contrast, Hyde and Linn (1988) concluded in a later meta-analysis that "the magnitude of the gender difference in verbal ability is currently so small that it can effectively be considered to be zero" (p. 64). These findings were supported in a later epidemiological study focused on the identification of specific language impairment in a diverse sample of kindergartners, which determined that boys are no more likely than girls to be identified as having a specific language impairment (Tomblin et al., 1997), debunking a long-held belief regarding significant differences in impairment rates for boys compared to girls.

Gender differences in reading

Gender differences in the development of important academic skills, including reading, are widely reported (Caro et al., 2009; Flannery et al., 2000; Kingdon, Serbin, & Stack, 2017; Liederman et al., 2005; McGeown et al., 2012; Raag et al., 2011). Similar to language, the gap in reading ability between boys and girls appears to widen with age (Logan & Johnston, 2010). Kingdon et al. (2017) examined the developmental reading trajectories of 126 low-income male and female children from elementary to secondary school and found that all children performed similarly in elementary school but that a gender gap in academic performance emerged in secondary school. That is, girls outperformed boys beginning in secondary school; boys continued to experienced difficulty through secondary school, while girls' performance remained stable.

In a meta-analysis focused on overall grades by gender across academic subject areas, Voyer and Voyer (2014) also found a consistent advantage for females from elementary school through the university level. The advantage was more pronounced for language-based subjects and less pronounced for math. Also similar to the gender advantage identified for language, the gender differences in reading in the Voyer and Voyer analysis evidenced relatively small effect sizes, which has been reported by others (McGeown et al., 2012). In contrast, in a study of low-income boys and girls, Kingdon et al. (2017) reported that both boys' and girls' performance decelerated in secondary school, with boys declining at a faster rate than girls. Furthermore, they found evidence of an advantage for females in both reading and math and estimated that females outperformed males by about 0.25 *SD* across all subject areas. This estimated effect size was reported for language as well (Maccoby & Jacklin, 1974).

It is important to note that many factors have been identified as contributing to the reading achievement gap between boys and girls, including motivation, differential attitudes toward reading, and differences in teacher and parent expectations of boys and girls. Overall, the literature in reading is more consistent than that in language. Whereas gender differences in language are debated, most investigators examining gender and reading agree that there is a difference in performance, with girls having an advantage over boys. The age at which the difference emerges and the magnitude of the difference continue to be discussed.

The language and reading of African American boys

Studies focused on the impact of gender on language and reading for African American children contrast significantly with the literature reported above, which either was not focused on or did not include African American children in substantial numbers. African American children overall perform poorly in school, and within this racial subgroup of American students, African American boys reportedly perform significantly more poorly than African American girls. Indeed, studies of achievement focused on African American children report large gender disparities beginning in infancy and persisting throughout schooling (Aratani et al., 2011; Flannery et al., 2000; Harris & Graves, 2010; Matthews et al., 2010; Roberts, Burchinal, & Durham, 1999).

Studies of the race and gender gap in education for African American children rely largely on data documenting the Black/White achievement gap reported nationally for fourth, eighth, and 12th graders on the NAEP, also called the *Nation's Report Card* (Bohrnstedt, Kitmitto, Ogut, Sherman, & Chan, 2015). Outcomes of the NAEP have documented a long-standing gap in achievement between African American children and their White and Asian peers in reading, mathematics, and other academic subject areas. When these race gap data are disaggregated further by gender there is a smaller measurable gap in performance between African American boys and both White and African American girls. In the current study, African American girls enrolled in the same schools and living in the same neighborhoods provided a natural comparison group.

Language

As reading is a consistent area of educational concern at the national, state, and local levels, there are many more studies focused on reading than on the general language use and development of African American boys. Studies of the language of African American students tend to focus more on language variation than on general language ability and seldom consider gender. Those language studies that are available compare African American boys either to White boys, describing a gap that is more race than gender based, or to African American girls. In a report focused on the race gap in very young children, Aratani et al. (2011) examined the performance of African American boys compared to White boys in preschool and kindergarten on a number of developmental measures, including language, utilizing a large sample from the Early Childhood Longitudinal Study–Birth Cohort. Across all experimental measures, including language and reading, White boys significantly outperformed African American boys. In the case of reading, African American boys scored 0.10 *SD* to 0.20 *SD* below their White peers in both preschool and kindergarten. The language gap was more significant at these ages and reportedly increased with age. For both reading and language, the race gap disappeared when SES was controlled, which suggests that perhaps the gaps observed are driven by low SES rather than race.

Roberts et al. (1999) examined the language skills of a longitudinal cohort of 87 low-income African American boys and girls at 18, 24, and 30 months of age. They measured vocabulary and grammatical development using both parent reporting and standardized language tests appropriate to the age of the participants. Results indicated that girls had larger vocabularies, used longer utterances, and used more irregular noun and verb forms than boys. Effect sizes were large for vocabulary differences at 24 months (Cohen's $d = 1.08$) and moderate for grammatical differences (Cohen's $d = 0.63$). The gender differences in grammatical development reported by Roberts et al. for low-income African American boys and girls were also found in an investigation of the narrative performance of older African American children (11–12.5 years; Mainess, Champion, & McCabe, 2012).

With the exception of Mainess et al. (2012), studies focused on the gender gap in language for African American boys mirror those for non-African American children in that they are focused on very young children. In the case of African American boys, however, the gaps in language have larger reported effect sizes, which suggests that African American girls perform significantly better than boys at these young ages regardless of SES. In addition, these outcomes are reportedly influenced by a host of sociodemographic and environmental variables that affect outcomes for boys, including SES, parental education, quality of the home environment, and amount of language input received (Aratani et al., 2011; Roberts et al., 1999; Tamis-LeMonda, Song, Leavell, Kahana-Kalman, & Yoshikawa, 2012).

Reading

Though still few in quantity, there have been several more papers focused on the poor reading achievement of African American boys than on general language skills, though many are not data based. It is important to note that many of these studies do not adequately address both race and gender. That is, the focus is often on the race gap between White and African American boys rather than the gender gap between boys and girls. For example, Fantuzzo et al. (2012) utilized a cumulative risk framework to examine the differences in performance of urban African American and White boys on a range of academic skills, including reading. As has been reported elsewhere, there was a reading gap between African American and White boys, with African American boys performing more poorly on reading assessments as well as on assessments of mathematics. This difference by race is well documented both in the extant literature (Husband, 2012a, 2012b; Matthews et al., 2010) and on the NAEP (Bohrnstedt et al., 2015).

In a longitudinal investigation focused on gender and race differences, Matthews et al. (2010) examined the literacy gap in a large sample of African American boys and girls and White boys and girls in kindergarten and first, third, and fifth grades from the Early Childhood Longitudinal Study–Kindergarten Cohort 1998–1999. Similar to many other studies, they found both gender and race gaps in the development of literacy in kindergarten between African American boys and their peers. The race gap was more pronounced than the gender gap, producing moderate effect sizes, but both gaps continued to increase in magnitude through fifth grade. These authors noted that the African American boys in the sample were also more likely to be from low-income homes with poor home literacy environments. The additive effects of gender and race were identified as influential in the poor performance of African American boys in the sample.

Utilizing the same kindergarten cohort and first-grade cohort, Chatterji (2006) reported similar outcomes for African American children in the early grades. They confirmed that the reading gaps between boys and girls and between African Americans and Whites continued to grow in size across these early grades and that these outcomes were most pronounced for children from low-income households, which characterized most of the African American boys in the sample.

McMillian, Frierson, and Campbell (2011) focused on the performance of a small ($n = 113$) sample of low-income African American boys and girls at ages 8 and 12, utilizing a secondary data analysis of children who participated in a randomized trial focused on educational intervention. They hypothesized that there would be no differences in mathematics or reading achievement at age 8 but that girls would outperform boys at age 12. However, the hypothesis was not supported, as no gender differences in either reading or mathematics achievement were evident at either age. Unfortunately, the sample size limited this study's statistical power to detect differences in a sample of children who exhibited very little variation in academic performance overall.

Overall, studies focused on the performance of African American boys on reading measures present mixed findings. Whereas many articles present an overwhelming impression that these boys are failing academically compared to their peers regardless of race (Davis, 2003; Husband, 2012a, 2012b), and data-based research supports the consistent presence of race differences in performance, gender differences are less clear. An exception is Justice, Invernizzi, Geller, Sullivan, and Welsch (2005), who found a gender gap between girls and boys but no race gap between African American and White boys in most early literacy skills. Taken together, these investigations suggest that African American boys may or may not perform differently from girls in terms of the development of reading skills. When these differences are present, however, the data suggest that they persist throughout schooling and that income status substantially influences performance.

Individual differences in academic trajectories

The current investigation applied an individual change score model to examine the reading and language trajectories of African American boys compared to girls in first through fifth grades. Though few studies were available in the literature to guide our thinking about language trajectories for African American boys, three distinct types of longitudinal findings have been reported in the literature for children who are at risk for failure as a result of a variety of sociodemographic variables (e.g., differences attributed to SES, school readiness or learning-related skills such as executive function and social skills; Matthews et al., 2010; McClelland, Acock, & Morrison, 2006). The first type of longitudinal finding is stable growth in reading or math achievement over time with no significant increases or decreases in performance across grades (Caro et al., 2009; McClelland et al., 2006; McMillian et al., 2011). The second type of longitudinal finding is a narrowing of performance differences over time between groups by race or gender (Curby, Rimm-Kaufman, & Ponitz, 2009; Iruka, Gardner-Neblett, Matthews, & Winn, 2014). The third type of finding is a widening of achievement gaps between groups over time (Caro et al., 2009; Chatterji, 2006; Curby et al., 2009; Matthews et al., 2010; McClelland et al., 2006).

Evidence from this literature indicates that the type of longitudinal trajectory may vary as a function of a number of variables, including age and grade. For example, Caro et al. (2009) examined math performance from first grade through high school as it related to the SES of participants. Math performance remained stable from second through sixth grades regardless of SES, but the gap between students from lower and higher SES backgrounds widened from seventh through 10th grades. Curby et al. (2009) found that children who demonstrated higher reading ability at the start of kindergarten grew faster than their peers who started out at lower levels, which implies that the gap between these two groups would widen in subsequent grades. Conversely, children who started out with higher performance in math and phonological awareness grew more slowly than their peers who started out lower, which suggests a narrowing of the gap in subsequent grades.

The current study

It is not clear from the evidence in the extant literature which of these three trajectories—stable growth over time, narrowing of performance differences, or widening of performance gaps—characterizes the reading and language growth of African American boys and how the trajectory is different compared to African American girls in the same schools. There have been both studies that report a widening of the gender gap throughout schooling (Aratani et al., 2011; Caro et al., 2009; Matthews et al., 2010; McClelland et al., 2006) and others that suggest that these differences may be resolved during the early years (Iruka et al., 2014). These studies further suggest that the trajectory may be quite different for language versus reading. The current longitudinal study permitted an examination of growth in both reading and language across a large sample of African

American boys and girls, all of whom attended the same schools. The performance of males was compared to that of females.

Results in the extant literature do not provide a clear picture of the gender differences in performance that may exist for African American boys. Is there an overarching boy problem, as has been suggested (Davis, 2003; Husband, 2012a, 2012b), such that the performance of African American boys is significantly different from that of girls? If differences are identified, how do they change with increasing age or grade? In order to address these important issues, we posed the following research question: How do African American boys and girls differ in the longitudinal development of reading, cognition, and language in first through fifth grades?

Method

Participants

Participants were enrolled in a larger project focused on language, literacy, and dialectal variation. African American boys and girls ($N = 890$) were investigated in first through fifth grades in a major urban school district in the southeastern United States. Participants ranged in age from 5.8 to 12.5 years of age ($M = 8.3$, $SD = 1.3$ years). Approximately half (48%) of the participants were male and half were female. At the beginning of each academic year, children were recruited for participation during school orientation sessions. Doctoral students and doctoral-level project personnel (e.g., research scientists and project coordinators) were responsible for attending the orientation to disseminate details of the study to families and distribute consent forms.

All children who returned consent forms were considered for inclusion in the current study. Participants attended seven different schools in the public elementary schools—four schools were traditional public schools and the remaining three were public charter schools. Across all seven schools, eligibility for the National School Lunch Program ranged from 50% to 100%, with the highest percentages of children eligible for free and reduced-price lunch attending the traditional public schools.

In the current study, we only included children who were not enrolled in special education services and had complete information on gender. Accordingly, 55 children receiving special education services and four children with missing information on gender were excluded from the current study, which resulted in a final sample of 831 children. The final sample was nearly evenly split by gender (girls = 437, boys = 394). These children had normal nonverbal intelligence ($M = 96.94$, $SD = 15.47$, on the Kaufman Brief Intelligence Test–Second Edition [KBIT]; Kaufman & Kaufman, 2004).

Assessment measures

Language

We measured children's language skills using three subtests of the Test of Language Development–Primary: 4th edition (TOLD–P:4; Hammill & Newcomer, 2008a) and the Test of Language Development–Intermediate: 4th edition (TOLD–I:4; Hammill & Newcomer, 2008b). The TOLD–P:4 was administered to participants 7 years and younger, and the TOLD–I:4 was administered to children 8 years of age and older. Accordingly, participants in first and second grades were administered the Picture Vocabulary, Syntactic Understanding, and Morphological Completion subtests of the TOLD–P:4, and participants in third, fourth, and fifth grades were administered Picture Vocabulary, Sentence Combining, and Morphological Comprehension. Children 8 years old and older who were unable to achieve basal on the particular subtests of the TOLD–I:4 were administered the corresponding subtests of the TOLD–P:4 (e.g., Morphological Completion was administered in place of Morphological Comprehension). These subtests assess children's receptive vocabulary, syntax, and morphological knowledge. Although the selected subtests of the two versions of the TOLD measure similar constructs, the scoring is based on different scales.

Unfortunately, the two versions of the TOLD are not vertically scaled—a score on the TOLD–P:4 cannot be compared in a meaningful mathematical way to a score on the TOLD–I:4, except via norm-referenced standard scores (called *scaled scores* in the manual, with a mean of 10 and standard deviation of 3). To overcome this lack of an appropriate developmental scale in our longitudinal sample, we fit a single-factor model of language to the second-grade students in the study. Second-grade students are at the recommended age boundary between the versions of the test (8 years old), and depending on their performance they were administered the TOLD–P:4 or TOLD–I:4.

Our resulting sample of second-grade students included 205 who took three TOLD–P:4 subtests, 110 students who took three TOLD–I:4 subtests, and 16 students who took a mixture of subtests of each version. Because the intention of the TOLD is to measure general language ability, we fit a confirmatory single-factor model to all six subtests for second grade that treated the test scores as missing at random and jointly scaled all tests to indicate latent language ability. This model fit excellently—$\chi^2 = 3.78$, $df = 6$, comparative fit index (CFI) = 1.00, Tucker–Lewis index (TLI) = 1.02, root mean square error of approximation (RMSEA) < 0.01—with good standardized loadings for the six subtests (median loading =0.73). We then applied the parameters (loadings and intercepts) of this second-grade model to the full sample of students across all grades for the versions of the tests they took. This was a model of strong invariance (Meredith, 1993; Vandenberg & Lance, 2000) that allowed latent variances and means to be estimated across grades. The resulting factor scores indicated a longitudinally consistent *z* score of latent language ability, using whatever subtests or versions participants took (estimated via full information maximum likelihood). This developmental *z* score was the language score used in the current longitudinal study, with the mean and variance centered on second-grade performance.

Reading

Five subtests of the Woodcock–Johnson III Tests of Achievement (Mather, 2001) were used to assess children's ability to process and manipulate phonological information. Reading Vocabulary, Word Attack, Reading Fluency, Passage Comprehension, and Letter-Word Identification measure literacy and language-related skills that support reading.

Reliability

We established reliability by double-entering and scoring approximately 20% of the data obtained for all assessment instruments. Establishing entry reliability as well as scoring reliability was designed to ensure the quality and accuracy of data input. We calculated scores electronically using scoring tables that were created based on raw scores from the published scoring criteria. Agreement for entry reliability was 96%, and scoring reliability was 100% for the sample.

Administration and scoring procedures

All data were collected by trained master's- and doctoral-level students in communication disorders and related fields (e.g., education, developmental psychology, counseling psychology) under the direction of the project coordinators and principal investigators. Data collectors were from various racial/ethnic backgrounds and were speakers of American English. Data were collected in each child's school in quiet spaces identified for use by school personnel. Examiners received training on each instrument as presented in the examiner's manuals prior to collecting data in

Table 1. Demographic characteristics by cohort (Wave 1).

			Age		Intelligence	
Study year	Grade	Male (%)	*M*	*SD*	*M*	*SD*
2013–2014	1	50	6.83	0.41	92.94	17.65
	2	40	7.88	0.49	96.51	15.17
	3	52	8.86	0.47	96.40	15.23
	4	48	9.96	0.57	95.32	13.58
2014–2015	1	45	6.86	0.48	97.56	14.70
	2	45	7.88	0.49	98.36	16.76
	3	47	8.77	0.46	102.82	14.82
	4	54	9.90	0.47	98.49	14.95
	5	58	11.33	0.60	97.75	9.26

Note. Intelligence is the standard score from the Kaufman Brief Intelligence Test. The table is given to describe the cohorts of the two-wave design, and intelligence and gender will be modeled longitudinally in subsequent results.

schools. Each instrument was scored by a second set of trained graduate students who were responsible for data reduction and who were supervised by project methodologists.

Research design

The larger project obtained 5-year longitudinal data from first- through fifth-grade African American students. An accelerated cohort design was used to test each participant across 2 years of the project. Specifically, in the first year of the project, we tested one cohort of participants, including 137 first graders, 134 second graders, 133 third graders, and 126 fourth graders; in the second year, we retested about 63% of these children, who were now second graders, third graders, fourth graders, and fifth graders. Meanwhile, in the second year, we also tested a new cohort of participants, including 127 first graders, 77 second graders, 90 third graders, 54 fourth graders, and 12 fifth graders; in the third year, we retested 41% of these children. Accordingly, 54% of the sample had two data points included in the data set. Children tested for the first time in fifth grade had only one data point. Table 1 shows the demographic characteristics of each cohort at each test occasion after we removed the 55 children involved in special education services and four children with missing gender information.

Data analysis

Although we were concerned with the academic performance of African American boys, the girls in the sample provided a natural comparison group, as they shared the same instruction and neighborhoods. Because this was a longitudinal sample, we used change score models to evaluate growth over time in Grades 1 to 5.

Change score models for growth

We used individual change score models to describe individual variability in intercept, proportional change, and constant change (the linear rate of change) in the form of dual change score models (McArdle, 2001; McArdle & Hamagami, 2001). We have used these models successfully to characterize the relationship between and growth in dialect variation and reading (Washington, Branum-Martin, Sun, & Lee-James, 2018). The general form of a dual change score model for a single outcome Y_{gi} for student i at grade g is as follows:

For $g = 1$,

$$Y_{1i} = Intercept_i + e_{1i} \quad (1)$$

For $g = 2$ to 5,

$$Y_{gi} = Y_{(g-1)i} + \Delta Y_{gi} + e_{gi} \tag{2}$$

$$\Delta Y_{gi} = \beta Y_{(g-1)i} + Slope_i, \tag{3}$$

where $Intercept_i$ is the student's individual starting point (i.e., the predicted initial score at Grade 1) and e_{1i} is random error (Eq. 1). Equation 2 shows how a student's score at any given time is a function of three things: the prior year's score ($Y_{(g-1)i}$), the amount of change (ΔY_{gi}), and random error (e_{gi}). Finally, Equation 3 describes the latent change score between the score at grade g (Y_{gi}) and the score at grade (g-1) for each individual, which concludes the individual linear rate of change ($Slope_i$; i.e., some students may grow faster than others) and the individual proportional change ($\beta Y_{(g-1)i}$), in which β is the proportional change parameter that describes the curvature of the growth trajectories. In addition, individual intercepts may covary with individual slopes.

The model is called *dual change* because of the linear growth combined with the proportional growth portion (Eq. 3). It can describe nonlinear trajectories with a proportional change parameter (fixed across individuals) that predicts status at any given time point as a function of the previous time point. Such a proportional change can be positive (accelerating growth, as in compound interest) or negative (decelerating growth, as in slowing to an asymptote). If the proportional change parameter happens to be zero, the change score model reduces to a model of individual linear growth.

The crucial test for the current study is that these growth models may differ between genders for all of the above parameters. Girls may start higher, grow faster, and have a different rate of proportional change—acceleration or deceleration may differ across genders. In addition, their variability may also differ: Girls may be more heterogeneous or homogeneous in intercepts and slopes than boys. Fortunately, these change score models can be readily tested across groups in a multiple-group structural equation modeling framework (Joreskog, 1970; Vandenberg & Lance, 2000; Little, 2013).

For each outcome, we present three models. First, we present an unrestricted (free) model for each of the two groups (i.e., a multiple-group structural equation modeling for growth). Second, because girls and boys might reasonably be expected to start schooling similarly, we present a model of equal initial status. Third, we present an omnibus model of no gender difference. We use this three-step process in order to isolate the nature of gender-based differences over time. If the third model of full equivalence fits, then there are no statistically dependable differences between boys and girls in their growth characteristics in this sample. If the third model fails but the second one fits, then girls and boys start equivalently but develop differently over time. Finally, if neither of these two restrictions fits, then girls and boys may differ in complex ways and should be evaluated separately, as if they each had their own growth model in Step 1.

Results

Table 2 presents descriptive statistics for boys and girls for each of the seven outcomes. Sample size, mean, standard deviation, skewness, and kurtosis are presented for each group. In addition, the two columns of difference present statistical tests of mean difference, with a t test and Cohen's d statistic for effect size. Most of the skewness and kurtosis values suggest that the performance did not deviate strongly from a normal shape. The rightmost two columns present the mean standard scores for each outcome for boys and girls. Five of the seven outcomes showed statistically significant ($p < .05$) mean differences favoring girls over boys. Regardless of statistical significance, the gender difference effect sizes for the reading outcomes ranged from 0.20 to 0.71 in Grades 4 to 5.

Table 2. Descriptive statistics.

	Boys					Girls					Difference		Standard score	
Grade	*N*	*M*	*SD*	Skewness	Kurtosis	*N*	*M*	*SD*	Skewness	Kurtosis	*t*	*d*	Boys	Girls
Kaufman Brief Intelligence Test, total score														
1	112	19.28	4.69	1.03	0.97	123	19.62	5.14	0.69	0.09	−0.53	.07	93.53	96.94
2	148	22.33	5.26	0.19	−0.85	184	22.65	5.56	0.20	−0.64	−0.53	.06	98.12	99.30
3	163	25.46	5.75	−0.20	−0.66	172	25.85	4.99	−0.15	−0.26	−0.67	.07	99.34	102.24
4	153	26.99	4.99	−0.29	−0.25	145	26.71	4.91	−0.44	−0.13	0.48	−.06	96.78	96.88
5	58	27.67	5.04	−0.02	−0.53	46	29.50	5.18	−0.44	1.36	−1.81	.36	92.56	97.27
Test of Language Development, factor score														
1	121	−0.55	0.95	−0.25	−0.10	127	−0.57	0.88	−0.05	−0.04	0.18	−.02	9.08	9.15
2	148	0.15	0.91	−0.02	−0.45	182	−0.05	0.97	0.16	−0.56	1.88	−.21	9.26	8.99
3	162	0.52	1.05	0.28	−0.52	172	0.32	1.11	0.36	−0.54	1.69	−.19	7.97	7.97
4	153	0.89	0.91	0.42	−0.29	144	0.86	1.00	0.18	−0.14	0.31	−.03	6.96	7.65
5	58	1.30	0.96	−0.07	−0.02	47	1.37	1.13	0.22	0.51	−0.33	.07	7.33	7.87
WJ Letter-Word Identification, *W* score														
1	122	430.75	30.48	0.49	0.30	128	435.31	28.20	0.14	0.49	−1.23	0.16	106.71	111.62
2	151	462.55	31.91	−0.28	−0.01	186	464.19	24.12	−0.02	−0.61	−0.52	0.06	103.01	104.65
3	167	483.40	25.25	−0.67	1.66	175	482.95	23.14	−0.99	2.03	0.17	−0.02	99.79	101.44
4	154	491.03	24.60	−0.41	0.55	145	495.50	18.62	−0.07	0.61	−1.78	0.20	96.90	99.89
5	59	498.32	24.19	−0.47	0.21	47	508.53	19.30	−0.50	−0.21	−2.36*	0.47	96.28	101.85
WJ Word Attack, *W* score														
1	122	465.87	25.65	−0.41	0.49	127	468.45	19.30	−0.76	2.54	0.89	0.11	109.02	113.39
2	151	480.97	21.33	−0.21	1.09	186	481.35	17.17	0.20	0.31	0.18	0.02	102.96	103.96
3	167	492.11	18.03	−0.13	0.71	174	489.56	16.77	−0.31	−0.04	−1.35	−0.15	100.80	100.48
4	154	492.62	19.72	−0.22	0.31	145	498.04	16.02	0.08	−0.44	2.61*	0.30	97.49	100.97
5	59	497.24	21.68	0.17	−0.61	47	505.83	16.38	0.29	−0.12	2.25*	0.44	98.36	102.98
WJ Reading Vocabulary, *W* score														
1	120	457.08	15.07	0.53	−0.73	128	460.92	16.16	0.56	−0.48	1.93	0.25	97.63	100.82
2	149	476.06	17.97	−0.13	−0.47	184	476.99	15.19	−0.13	−0.34	0.51	0.06	99.41	100.27
3	164	487.04	16.01	−0.15	0.11	171	486.55	14.49	−0.44	0.79	−0.29	−0.03	98.47	99.93
4	153	490.37	15.11	−0.20	0.65	145	493.63	13.78	0.07	0.47	1.94	0.23	95.37	98.03
5	57	495.51	11.67	−0.34	1.78	47	503.57	13.05	0.68	0.56	3.32*	0.66	92.58	101.34
WJ Passage Comprehension, *W* score														
1	122	450.80	21.77	−0.41	−0.49	128	455.48	18.68	−0.11	−0.26	−1.82	0.23	99.14	104.10
2	151	472.41	16.82	−0.57	0.21	186	473.56	13.13	−0.23	0.00	−0.69	0.08	97.54	98.87
3	167	481.16	12.84	−2.03	13.57	175	480.37	11.88	−0.99	3.10	0.60	−0.06	93.44	94.27
4	154	483.62	11.83	−0.46	0.73	145	485.74	9.66	−0.17	0.18	−1.70	0.20	88.97	91.54
5	59	487.20	9.32	−0.37	−0.03	47	493.40	9.25	−0.17	−0.52	−3.41*	0.67	86.74	92.32
WJ Reading Fluency, *W* score														
1	86	446.41	18.86	0.00	0.09	93	449.62	17.55	0.81	3.95	1.18	0.18	106.53	109.55
2	139	462.87	19.70	−0.05	0.74	180	464.54	16.30	0.02	1.05	0.81	0.09	100.31	102.33
3	162	474.89	16.82	0.50	0.05	168	476.41	14.33	−0.08	2.52	0.88	0.10	96.26	98.90
4	152	483.72	19.32	0.69	1.10	145	489.83	21.12	0.70	0.62	2.60*	0.30	93.84	97.79
5	57	490.44	19.98	0.57	−0.14	47	508.11	30.33	1.62	3.72	3.43*	0.71	91.44	99.46

Note. The expected standard score is 10 for the Test of Language Development and 100 for the other six outcomes. WJ = Woodcock–Johnson III Tests of Achievement.
*$p < .05$, differences are scaled in favor of girls.

Growth model tests

Because these were longitudinal data, and many students appeared in two grades, these grade-wise differences can be better evaluated for what they suggest about growth over time. Table 3 presents the results of the three growth models for each of the seven outcomes. Table 3 presents conventional fit statistics, including chi-square, longitudinal CFI (Little, 2013), TLI, and RMSEA. Most models fit reasonably well overall (Little, 2013; Marsh, Hau, & Grayson, 2005; Marsh, Hau, & Wen, 2004) and are discussed in detail. The constrained models for each outcome also had nested model comparisons, including a *p* value for the chi-square difference (p[LRT]), the change in CFI (Cheung & Rensvold, 2002; Little, 2013; Marsh et al., 2005; Marsh et al., 2004), and the final column of Table 3 noting whether the model suggested gender equality for the parameters being tested.

Table 3. Change score model tests of equality across genders.

Outcome	Model	χ^2	*df*	CFI	TLI	RMSEA	p(LRT)	ΔCFI	Gender equality
Kaufman Brief Intelligence Test	Free[a]	62.24	18	.909	.878	.078			
	Intercept equal[a]	62.25	20	.913	.895	.072	.995	.004	Yes
	All equal[a]	63.18	22	.915	.907	.068	.629	.002	Yes
Test of Language Development	Free[a]	26.36	18	.988	.984	.034			
	Intercept equal[a]	27.69	20	.989	.987	.031	.512	.001	Yes
	All equal[a]	31.53	22	.986	.985	.033	.147	−.003	Yes
WJ Letter-Word Identification	Free	54.66	14	.972	.952	.084			
	Intercept equal	56.49	16	.972	.958	.078	.402	.000	Yes
	All equal	70.59	20	.965	.958	.078	.007	−.007	Yes
WJ Word Attack	Free	40.59	14	.967	.944	.068			
	Intercept equal	49.90	16	.958	.937	.072	.009	−.009	Yes
	All equal	57.88	20	.953	.944	.068	.092	−.005	Yes
WJ Reading Vocabulary	Free[a]	102.31	18	.927	.903	.107			
	Intercept equal[a]	103.20	20	.928	.914	.101	.642	.001	Yes
	All equal[a]	116.69	22	.918	.911	.103	.001	−.010	Yes
WJ Passage Comprehension	Free[a]	106.35	18	.920	.894	.109			
	Intercept equal[a]	114.29	20	.915	.898	.107	.019	−.005	Yes
	All equal[a]	128.58	22	.904	.895	.109	.001	−.011	Not equal
WJ Reading Fluency	Free	34.03	14	.977	.961	.061			
	Intercept equal	39.69	16	.973	.959	.062	.059	−.004	Yes
	All equal	62.84	20	.951	.941	.075	.000	−.022	Not equal

Note. See the text for a discussion of model fit. CFI = comparative fit index, evaluated against a longitudinal null model (Little,2013); TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; p(LRT) = *p* value from the likelihood ratio test as a difference in chi-square versus the free model; ΔCFI = change in CFI versus the free model; WJ = Woodcock–Johnson III Tests of Achievement.

[a]Slope variance fixed at zero to achieve convergence for both genders.

For the tests of model restriction, the *p* value for the likelihood ratio test (p[LRT]) notes whether the model with parameters constrained to be equal across genders fits just as well as the freely estimated two-group model (i.e., $p > .05$ would suggest equality across genders). However, the LRT is noted to be rather conservative, rejecting constrained models that might otherwise be reasonable (Cheung & Rensvold, 2002). Therefore, we also report the change in CFI, in which a decrease greater than .010 indicates substantial misfit for the constraints or that there is a gender difference in the constrained parameters.

The free model had reasonable fit for most outcomes (CFI, TLI > .90, RMSEA < .10), which indicates that a dual change score model fit jointly to both genders was a reasonable characterization of the growth trajectories. Although there are no objective criteria for evaluating the fit of these longitudinal models (Little, 2013; Marsh et al., 2005; Marsh et al., 2004), the free model for KBIT and passage comprehension showed some misfit on TLI (less than .90). In addition, the free model for passage comprehension and reading vocabulary also showed some misfit, with RMSEAs of .109 and .107, respectively.

The first model restriction (Step 2) was to test for equal starting points across genders in first grade. The intercepts-equal model fit in an absolute sense (p[LRT] < .05) for five of the seven outcomes. The more reasonable test of ΔCFI fit for all seven outcomes (ΔCFI < −.010), which suggests that in the context of a longitudinal model of all five grades on average boys and girls perform similarly in first grade. The second model restriction, the omnibus test of total equality across genders (Step 3), fit in an absolute sense (p[LRT] < .05) for four of the seven outcomes. The more reasonable test of ΔCFI suggests that growth trajectories are completely equal across genders for all outcomes except passage comprehension and reading fluency (ΔCFI < −.010).

Model results

The resulting parameters of these dual change score models (Eqs. 1–3) are presented in Tables 4 and 5. Table 4 presents the parameter estimates for the five outcomes that resulted in equality

Table 4. Model estimates for outcomes with equal growth parameters across genders.

Outcome	Parameter	Estimate	*SE*
Kaufman Brief Intelligence Test	Intercept mean	19.95	0.30
	Slope mean	3.64	0.64
	Proportional change	−0.06	0.03
	Intercept variance	18.25	2.33
	Slope variance	—	—
	Covariance (intercept, slope)	—	—
	Residual	12.43/11.64	1.21/1.06
Test of Language Development	Intercept mean	−0.55	0.05
	Slope mean	0.47	0.02
	Proportional change	0.03	0.02
	Intercept variance	0.63	0.06
	Slope variance	—	—
	Covariance (intercept, slope)	—	—
	Residual	0.24/0.23	0.02/0.02
WJ Letter-Word Identification	Intercept mean	21.77	0.08
	Slope mean	8.86	0.51
	Proportional change	−0.34	0.02
	Intercept variance	2.27	0.19
	Slope variance	0.17	0.03
	Covariance (intercept, slope)	0.25	0.06
	Residual	0.12/0.10	0.02/0.02
WJ Word Attack	Intercept mean	23.41	0.06
	Slope mean	8.16	1.13
	Proportional change	−0.32	0.05
	Intercept variance	1.03	0.10
	Slope variance	0.10	0.03
	Covariance (intercept, slope)	0.12	0.05
	Residual	0.18/0.11	0.02/0.02
WJ Reading Vocabulary	Intercept mean	23.09	0.05
	Slope mean	3.60	0.39
	Proportional change	−0.13	0.02
	Intercept variance	0.72	0.06
	Slope variance	—	—
	Covariance between intercept and slope	—	—
	Residual	0.11/0.11	0.01/0.01

Note. Estimates before the slash are for boys, and those after the slash are for girls (residual variances and their standard errors). Dashes indicate a parameter constrained to zero. WJ = Woodcock–Johnson III Tests of Achievement.

across genders (residual variances were allowed to differ; Little, 2013), including the model-predicted mean for intercept (status at first grade), slope (linear change), and proportional change. The variances of the intercept, slope, and residuals are given, along with the covariance, if applicable.

The first outcome in Table 4 shows the model estimates for intelligence (KBIT), with a total score of 19.95 items and a mean linear rate of change of 3.64 items per year. The proportional change was −0.06, which indicates that growth decelerated by 6% per year. The intercept variance of 18.25 suggests that there was wide variability ($SD = 4.3$ items) in first grade. There was no estimable variance in slope, which suggests that students grew in parallel trajectories. The TOLD was estimated on a factor score (z scale) centered at second grade. The intercept mean of −0.55, combined with mean slope of 0.47, suggests that first graders started about 0.5 *SD* below second graders ($z \approx 0$) and grew an average of nearly 0.5 *SD* per year. Proportional change was fairly close to zero (3%), which suggests mostly linear change.

The three Woodcock–Johnson tests in Table 4 were on a *W* score, but in order to get model convergence, we rescaled them by dividing by 20 points (this linear transformation did not change statistical features of the model or substantive interpretation). Letter-word identification and word attack had similar intercept, slope, and proportional change estimates (−0.34 and −0.32, respectively, suggesting decelerating growth). Both had estimated slope variances, and the covariance indicated a positive relation between intercept and slope ($r = .39$ for letter-word

Table 5. Model parameter estimates for outcomes with unequal growth parameters across genders.

		Boys		Girls	
Outcome	Parameter	Estimate	*SE*	Estimate	*SE*
WJ Passage Comprehension	Proportional change	−0.27	0.02	−0.34	0.02
	Intercept mean	22.80	0.05	22.80	0.05
	Slope mean	6.88	0.53	8.38	0.50
	Intercept variance	1.04	0.09	1.04	0.09
	Slope variance	—	—	—	—
	Covariance (intercept, slope)	—	—	—	—
	Residual	0.14	0.02	0.12	0.01
WJ Reading Fluency	Proportional change	−0.14	0.06	0.00	0.05
	Intercept mean	22.40	0.06	22.40	0.06
	Slope mean	3.89	1.32	0.66	1.25
	Intercept variance	0.72	0.10	0.72	0.10
	Slope variance	0.07	0.03	0.17	0.03
	Covariance (intercept, slope)	0.00	0.07	−0.25	0.06
	Residual	0.19	0.03	0.19	0.03

Note. Dashes indicate a parameter constrained to zero. WJ = Woodcock–Johnson III Tests of Achievement.

identification and $r = .37$ for word attack), which suggests that students who started higher grew faster. Reading vocabulary in Table 4 had a mean intercept of 23.09 ($SD = 0.84$) with a mean slope of 3.6 units and a proportional change of −13%. There was no estimable variance in slopes.

Table 5 presents results for the dual change score models for the two outcomes that had gender differences: passage comprehension and reading fluency. The layout of the table is similar to that of Table 4 but with separate columns for each gender. For passage comprehension, the mean slope for boys was 6.88 units, whereas for girls it was 8.38. The proportional change parameters were also fairly different: −27% for boys and −34% for girls. The results for reading fluency in Table 5 suggest strong difference in growth. The mean slope was 3.89 units for boys and 0.66 for girls. The proportional change parameter was −.14 for boys and almost zero (.003) for girls.

Visualizing the model implications

Because these dual change score models are complex, they are best interpreted visually—especially for the models in which multiple parameters differ across groups. Figures 1–3 show growth plots for each of the five outcomes that did not show gender differences. Figure 1 shows KBIT (top) and language (TOLD; at bottom), each with a gray line for each student in the two-wave accelerated cohort design. Each plot also has a dark line for the predicted average from the dual change score model. These graphs are helpful in that the simultaneous effects of the linear and proportional change parameters are shown. Figure 1 shows that average growth in both intelligence and language appears essentially linear in Grades 1 to 5.

Figure 2 presents individual student results (gray) and model-based predictions for letter-word identification (top) and word attack (bottom). Both panels present a slight deceleration over time. Figure 3 presents student data and model predictions for reading vocabulary, also showing slight deceleration but no difference across genders. Figure 4 presents the growth plots for the two outcomes that showed gender differences. Passage comprehension at the top of Figure 4 shows girls growing slightly faster than boys, with both having a deceleration. Reading fluency at the bottom of Figure 4 shows a high degree of similarity in the early grades, but while the girls grow essentially linearly, boys have substantial deceleration, evident by fifth grade.

Discussion

Using a longitudinal, accelerated cohort design, we examined the performance of African American children on language and reading measures across first through fifth grades. Overall,

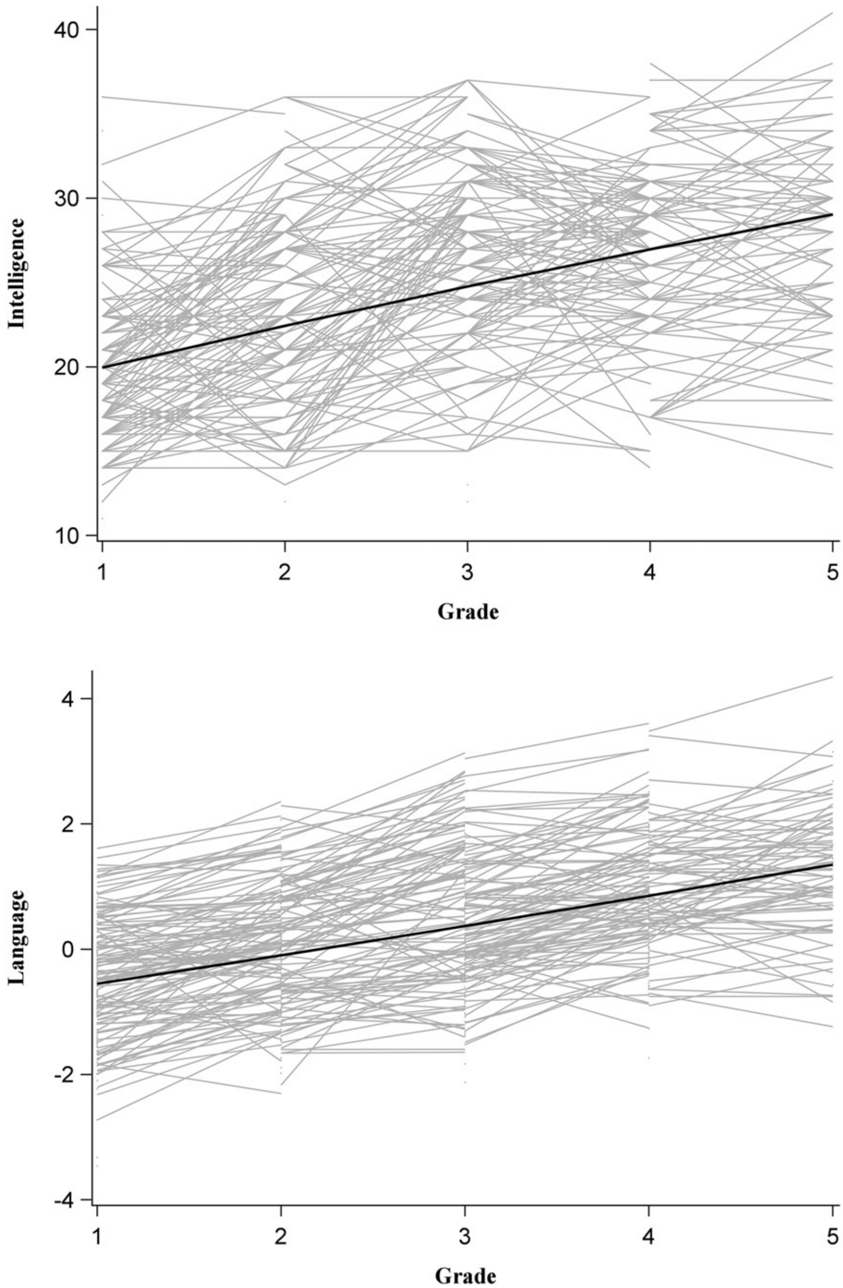


Figure 1. Individual growth plot with model-predicted change line for intelligence (top) and language (bottom).

the outcomes demonstrated that there was no evidence of gender differences in language or cognition in first through fifth grades. For reading comprehension and fluency, boys and girls performed equally in the early grades (i.e., first through third grades), but differences by gender emerged in fourth and/or fifth grade. Findings for language suggested that boys and girls performed equally. There were no statistically significant performance differences by gender. Furthermore, growth models indicated that African American boys and girls evidenced similar growth trajectories for language in first through fifth grades as well. These findings support the results of an early meta-analysis by Hyde and Linn (1988) that demonstrated that even when

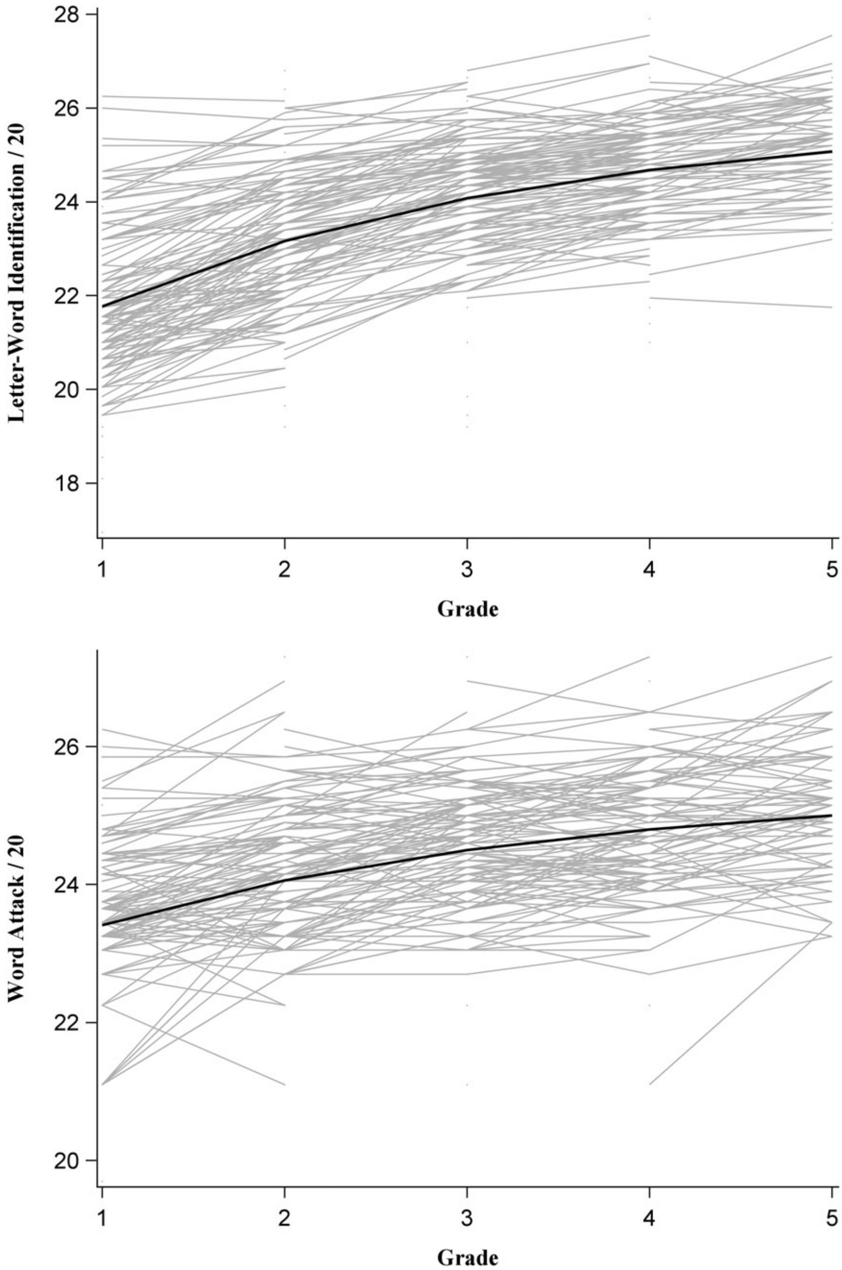


Figure 2. Individual growth plot with model-predicted change line for WJ Letter-Word Identification (top) and WJ Word Attack (bottom). WJ = Woodcock–Johnson III Tests of Achievement.

gender differences were apparent in language skills, the magnitude of the difference was so small as to be considered insignificant or nonexistent.

The results of the current investigation support the statement that these differences simply do not exist. Maccoby and Jacklin (1974), in contrast, performed a meta-analysis of 85 articles focused on gender and language and reported a pattern such that girls and boys were similar or equal early on language skills, but in preadolescence (∼10 years of age) and beyond these skills diverged, with significant deceleration evident for boys. That pattern was not identified in the current sample of African American boys and girls for language, but it was apparent for reading.

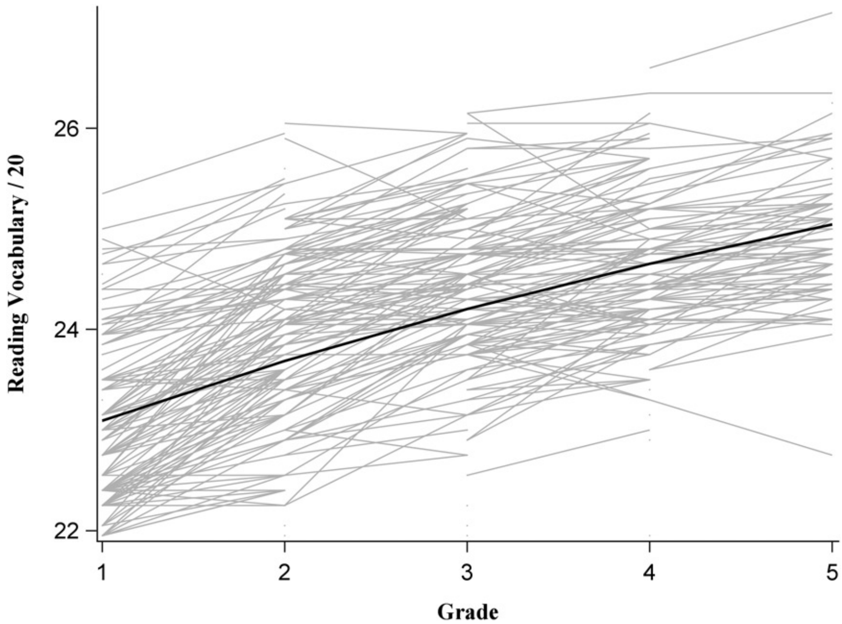


Figure 3. Individual growth plot with model-predicted change line for Woodcock–Johnson III Tests of Achievement Reading Vocabulary.

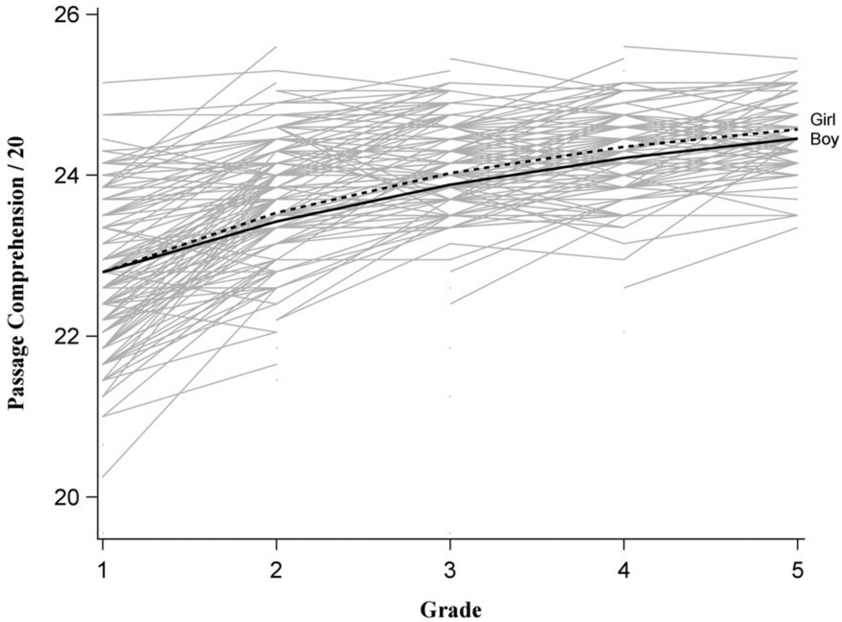
Instead, our outcome for language supports the 70% of studies reviewed by Maccoby and Jacklin (1974) that showed no differences in linguistic development for boys and girls.

It is important to note that the children in the current investigation were older than those included in many studies focused on gender differences in language. Studies with younger participants report that gender differences are apparent in infancy, toddlerhood, and preschool but disappear by kindergarten (Aratani et al., 2011; Bauer et al., 2002; Eriksson et al., 2012; Iruka et al., 2014; Maccoby & D'Andrade, 1966). Perhaps that was true for the participants in this investigation who showed no evidence of gender differences in language performance or growth by first grade.

Our findings contrast with the small number of studies focused specifically on gender differences in the language skills of African American children (Mainess et al., 2012; Roberts et al., 1999). These investigations, which were focused on preschoolers (Roberts et al., 1999) and preadolescents (Mainess et al., 2012), suggest that the language of African American girls is more varied and complex than that of African American boys and that these differences persist (Aratani et al., 2011). All of these investigations were focused on the performance of boys (and girls) from low-income homes. The current sample included largely low-income participants as well.

The Mainess et al. (2012) and Roberts et al. (1999) investigations both had small sample sizes; thus, it is possible that the variability in the language performance of their participants led to findings that would not be replicated with a larger sample with more power to detect both similarities and differences in performance. Perhaps important as well, though the current investigation and these earlier ones all examined gender differences in language, the language skills targeted were different. It is possible that with language, as is true with reading, gender differences can exist in the development of specific language skills or domains but not others (Tomblin et al., 1997). Aratani et al. (2011), though focused on African American boys, studied the race gap in achievement, which has a strong evidence base that suggests that these differences persist in all domains when African American boys are compared to their White and Asian peers.

Reading skill performance and growth in selected reading skills differed by gender in that reading trajectories diverged over time. No gender differences were apparent for any of the five



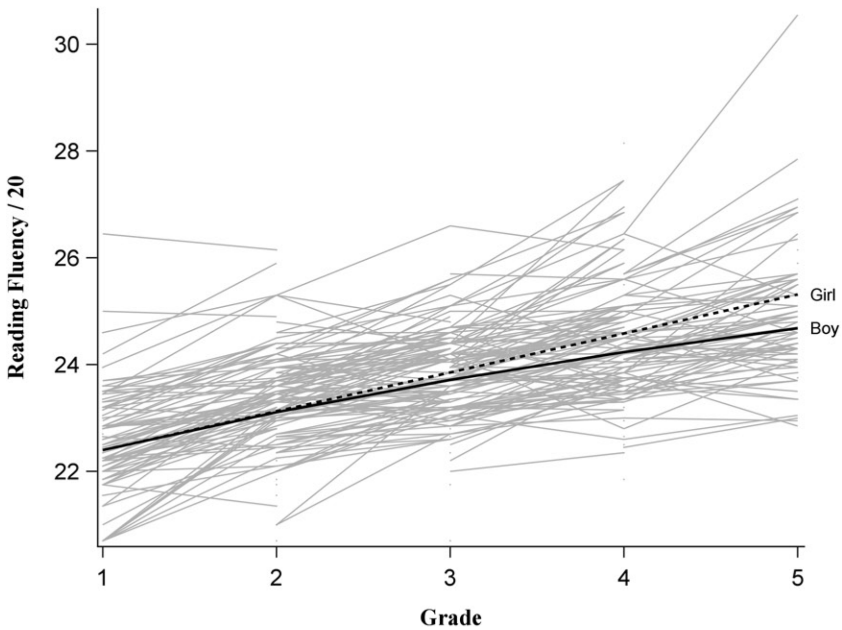


Figure 4. Individual growth plot with model-predicted change line for WJ Passage Comprehension (top) and WJ Reading Fluency (bottom). WJ = Woodcock–Johnson III Tests of Achievement.

reading skills measured in Grades 1–3. Gender differences in the descriptive statistics were not always consistent across these skills in Grades 4 and 5. Specifically, statistically significant differences in favor of girls were apparent in Grades 4 and 5 for reading fluency and word attack and in Grade 5 only for letter-word identification, passage comprehension, and reading vocabulary. Several other investigations have reported that the reading gap between boys and girls becomes apparent with advancing grade. Several have identified secondary school as the time at which this gap becomes apparent (Caro et al., 2009; Kingdon et al., 2017; Logan & Johnston, 2010; Voyer & Voyer, 2014). The growth models used suggest that not all of these observed differences are dependable, as girls differed from boys only on reading comprehension and reading fluency. Further longitudinal research could help to clarify which of these observed differences persist, grow, or disappear.

The current investigation documented gender differences in reading abilities that emerged earlier than secondary school; the boys' reading performance began to diverge as early as fourth grade. It is important to note that several of the skills on which African American boys showed decline in both fourth and fifth grades were skills that have been identified as early literacy skills that are foundational to becoming a good reader in later grades (Scarborough, 2001). Specifically, word attack and letter-word identification are critical for reading at the word level. As a result, they are both often the intense focus of early literacy programs. Perhaps most important, these skills support the development of more advanced reading skills, such as comprehension and fluency (Archer, Gleason, & Vachon, 2003; Cutting & Scarborough, 2006; Scarborough, 2001). Accordingly, weaknesses in these key early literacy skills would likely not provide a strong enough reading foundation to support the more complex reading required in fourth and fifth grades, in which content vocabulary is greater and the reading material becomes more syntactically complex and abstract. Thus, though boys performed only slightly below the mean on the TOLD and there were no gender differences, their early literacy skills appear not to be sufficiently strong to support the development of later reading abilities such as comprehension and fluency that rely on these early language skills.

Individual growth modeling also showed that the African American boys in this investigation experienced consistent growth in reading skills in the early grades, with divergence from girls occurring in the upper grades. Growth trajectories were essentially equal across genders for all outcomes except passage comprehension and reading fluency. Specifically, on the passage comprehension task, girls grew slightly faster than boys, but both boys and girls showed evidence of deceleration. With respect to reading fluency, girls were observed to grow linearly, whereas boys had substantial deceleration in fifth grade. This pattern has been described in the literature, whereby the reading trajectory for girls progresses more quickly and the trajectory for boys slows rapidly with age. Research suggests that this is the beginning of a gap in performance by gender (and SES) that will continue to widen with increasing age and grade (Kingdon et al., 2017; Logan & Johnston, 2010).

There are perhaps other variables that influence these outcomes for African American boys that were not considered in this investigation. Several, including motivation, the development of learning-related skills (e.g., executive function and social skills), differential interest in reading, gender-relevant curriculum content, and externalizing behaviors, among others, have been implicated (Bristol, 2015; Davis, 2003; Husband, 2012b; Matthews et al., 2010; McClelland et al., 2006). A likely explanation based on the trends identified in this study is that these boys simply have developed weak overall reading skills that catch up with them in later grades; reading comprehension and fluency become casualties of these weaknesses. In order to become skilled readers students must be able to connect their early literacy skills to later reading ability (Scarborough, 2001).

Scarborough (2001) divided reading skills into strands that early readers must master. According to Scarborough, most children who struggle with reading exhibit difficulty with mastery of the word recognition strand, which includes skills like phonemic and phonological awareness and letter-word identification. This pattern of reading difficulty characterizes the performance of the African American boys in the current investigation. Scarborough further suggested that later reading difficulty is often related to weaknesses in preschool language skills, even when these skills appear to resolve in later grades. This is particularly true for children for whom there is a family history of reading difficulty. In the current investigation, boys and girls performed and grew equally on measures of syntax, semantics, and morphology, which suggests that language differences did not differentiate the children in this sample, though this may have been true at younger ages. Furthermore, growth in language and reading related to the use of African American English dialect was examined using this same sample of 831 children (Washington et al, 2018). Though dialect use impacted reading, no gender differences were observed.

Why are these differences apparent for African American boys but not girls? This is the question that we all seek to answer. School-related variables are often used to explain these differences. In particular, quality of the school environment and teacher quality are often implicated in the poor performance of boys. However, the boys and girls in this investigation were recruited from the same schools, neighborhoods, and classrooms and were exposed to similar teaching and classroom environments. The SES backgrounds of the students were also similar. Perhaps it is some of the social variables discussed by others at the child level related to issues such as motivation and interest in reading, curriculum content, or contextual variables related to disciplinary practices and expectations that could differ across genders.

These variables may be explanatory to some extent, but they were not examined in the current investigation and thus are beyond our current scope. What is clear is that the pattern of gender differences and growth for African American boys is very similar to that reported in the larger literature for non-African American students whose reading achievement is driven by mastery of word-level abilities. Furthermore, the pattern of reading skill weaknesses identified for African American boys mirrors those identified for most children who struggle with reading, as indicated by Scarborough (2001). These outcomes suggest that a reading instructional strategy that focuses specifically on these foundational literacy skills may be needed to improve the outlook for this population of students.

Conclusions and limitations

The purpose of this investigation was to characterize growth in and the development of reading and language skills among African American boys and to identify any existing gender differences in their performance compared to girls. Boys and girls performed similarly on language and intelligence tasks, but there were clear differences in reading-related measures always favoring girls. Growth trajectories for reading also differentiated boys and girls in fifth grade. An important overall finding is that the patterns of language and reading growth observed for African American boys were similar to the patterns reported for the general population of students who struggle with reading, which suggests that improving access to current interventions utilized with these students should be beneficial for African American boys as well. Research suggests that the deceleration of reading skills identified among African American boys at fourth and fifth grades will continue. Future research should focus on reading and language growth beyond elementary school. Though there may be underlying social and motivational concerns contributing to these outcomes, as has been suggested in the extant literature, future research focused on the development of strong reading pedagogy and interventions would benefit these boys.

A limitation of the current study is that the primary aims were to characterize reading, so only one measure of cognition and one measure of language were used. Although reading was well covered by multiple measures, future research may help substantiate the current findings by using more varied measures of cognition and language. Multivariate models relating change across outcomes and their mutual, longitudinal influence may also be informative. Another limitation of the current investigation is that performance by classroom was not considered. Although six of the seven schools in the current sample were highly similar to one another in terms of racial makeup (>90% of students were African American) and low-income status (>90% of students qualified for free or reduced-price lunch), differences across classrooms may explain some of the variability. However, because students change classrooms yearly (are cross-classified), it is unclear how such variability would contribute to the shape of dual change score models (and most software cannot handle such cross-classified nesting in a multivariate 5-year model).

Although the accelerated cohort design yields 5 years of longitudinal data, no child was measured on more than two occasions. The overall shape of the trajectories may be adequately captured, but variability and individual differences in shape are poorly estimated. Future research with longer term designs and more measurement points may better explain individual variation in trajectories and provide stronger across-group comparisons (see Caro et al., 2009, for a cogent discussion of this point). Finally, the current sample ended at Grade 5. Further research would be helpful to determine whether the current trends stabilize, diverge, or disappear in later grades.

ORCID

Congying Sun http://orcid.org/0000-0002-5241-4462

References

Anastasi, A. (1958). Heredity, environment, and the question how? *Psychological Review*, *65*(4), 197.

Aratani, Y., Wight, V. R., & Cooper, J. L. (2011). Racial Gaps in Early Childhood: Socio-emotional Health, Developmental, and Educational Outcomes Among African-American Boys. Retrieved from Malman School of Public Health:

Archer, A. L., Gleason, M. M., & Vachon, V. L. (2003). Decoding and fluency: Foundation skills for struggling older readers. *Learning Disability Quarterly*, *26*(2), 89–101.

Bauer, D. J., Goldfield, B. A., & Reznick, J. S. (2002). Alternative approaches to analyzing individual differences in the rate of early vocabulary development. *Applied Psycholinguistics*, *23*, 313–335.

Bohrnstedt, G., Kitmitto, S., Ogut, B., Sherman, D., & Chan, D. (2015). School Composition and the Black-White Achievement Gap. Retrieved from Washington, DC: http://nces.ed.gov/pubsearch

Bristol, T. J. (2015). Teaching boys: Towards a theory of gender-relevant pedagogy. *Gender and Education*, *27*(1), 53–68.

Caro, D. H., McDonald, J. T., & Willms, J. D. (2009). Socioeconomic status and academic achievement trajectories from childhood to adolescence. *Canadian Journal of Education*, *32*(3), 558–590.

Chatterji, M. (2006). Reading Achievement Gaps, Correlates, and Moderators of Early Reading Achievement: Evidence From the Early Childhood Longitudinal Study (ECLS) Kindergarten to First Grade Sample. *Journal of Educational Psychology*, *98*(3), 489–507.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255.

Connell, J. P., Spencer, M. B., & Aber, J. L. (1994). Educational risk and resilience in African-American youth: Context, self, action, and outcomes in school. *Child development*, *65*(2), 493–506.

Curby, T. W., Rimm-Kaufman, S. E., & Ponitz, C. C. (2009). Teacher–child interactions and children's achievement trajectories across kindergarten and first grade. *Journal of Educational Psychology*, *101*(4), 912.

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, *10*(3), 277–299.

Davis, J. E. (2003). Early schooling and academic achievement of African American males. *Urban Education*, *38*(5), 515–537.

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., … Gallego, C. (2012). Differences between girls and boys in emerging language skills: evidence from 10 language communities. *British Journal of Developmental Psychology*, *30*(2), 326–343. doi:10.1111/j.2044-835X.2011.02042.x

Fantuzzo, J., LeBoeuf, W., Rouse, H., & Chen, C.-C. (2012). Academic achievement of African American boys: A city-wide, community-based investigation of risk and resilience. *Journal of School Psychology*, *50*(5), 559–579. doi:10.1016/j.jsp.2012.04.004

Flannery, K. A., Liederman, J., Daly, L., & Schultz, J. (2000). Male prevalence for reading disability is found in a large sample of black and white children free from ascertainment bias. *Journal of the International Neuropsychological Society*, *6*(4), 433–442.

Hammill, D. D., & Newcomer, P. L. (2008a). *Test of language development-primary*, 4th Edition. Austin, TX: Pro-Ed.

Hammill, D. D., & Newcomer, P. L. (2008b). *Test of language development – intermediate*, 4th Edition. Austin, TX: Pro-Ed.
Harris, T. S., & Graves, S. L. Jr, (2010). The influence of cultural capital transmission on reading achievement in African American fifth grade boys. *Journal of Negro Education*, *79*(4), 447–457.
Husband, T. (2012a). Addressing reading underachievement in African American boys through a Multi-Contextual approach. *Reading Horizons*, *52*(1), 1–25.
Husband, T. (2012b). Why can't Jamal read? *Phi Delta Kappan*, *93*(5), 23–27.
Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236–248.
Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin*, *104*(1), 53–69. doi:10.1037/0033-2909.104.1.53
Iruka, I. U. (2017). Perception is not always reality. *Exchange (19460406)*, *39*(234), 46–47.
Iruka, I. U., Gardner-Neblett, N., Matthews, J. S., & Winn, D.-M. C. (2014). Preschool to kindergarten transition patterns for African American boys. *Early Childhood Research Quarterly*, *29*(2), 106–117. doi:10.1016/j.ecresq.2013.11.004
Jencks, C., & Phillips, M. (2011). *The Black-White test score gap*. Washington, DC: Brookings Institution Press.
Jöreskog, K. G. (1970). A general method for estimating a linear structural equation system. *ETS Research Bulletin Series*, *1970*(2), i-41.
Justice, L. M., Invernizzi, M., Geller, K., Sullivan, A. K., & Welsch, J. (2005). Descriptive-developmental performance of at-risk preschoolers on early literacy tasks. *Reading Psychology*, *26*(1), 1–25.
Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test* 2nd Edition. Bloomington, MN: Pearson.
Kidd, E., & Lum, J. A. G. (2008). Sex differences in past tense overregularization. *Developmental Science*, *11*(6), 882–889. doi:10.1111/j.1467-7687.2008.00744.x
Kingdon, D., Serbin, L. A., & Stack, D. M. (2017). Understanding the gender gap in school performance among low-income children: A developmental trajectory analysis. *International Journal of Behavioral Development*, *41*(2), 265–274.
Liederman, J., Kantrowitz, L., & Flannery, K. (2005). Male vulnerability to reading disability is not likely to be a myth: a call for new data. *Journal of Learning Disabilities*, *38*(2), 109–129.
Little, T. D. (2013). *Longitudinal structural equation modeling*. New York, NY: Guilford Press.
Logan, S., & Johnston, R. (2010). Investigating gender differences in reading. *Educational Review*, *62*(2), 175–187.
Maccoby, E. E., & D'Andrade, R. G. (1966). *The development of sex differences* (*Vol. 5*). Redwood City, CA: Stanford University Press.
Maccoby, E. E., & Jacklin, C. N. (1974). Myth, reality and shades of gray: What we know and don't know about sex differences. *Psychology Today*, *8*(7), 109–112.
Mainess, K. J., Champion, T. B., & McCabe, A. (2002). Narration by African American Preadolescents - Preliminary Examination of Gender and Socioeconomic Issues. *Linguistics and Education*, *13*(2), 151–173.
Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit in structural equation models. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Contemporary psychometrics: A festschrift for Roderick P. McDonald* (pp. 275–340). Mahwah, NJ: Lawrence Erlbaum.
Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(3), 320–341. doi:10.1207/s15328007sem1103_2
Mather, N. (2001). Woodcock-Johnson III tests of achievement: Examiner's manual: Riverside Pub.
Matthews, J. S., Kizzie, K. T., Rowley, S. J., & Cortina, K. (2010). African Americans and Boys: Understanding the Literacy Gap, Tracing Academic Trajectories and Evaluation the Role of Learning-Related Skills. *Journal of Educational Psychology*, *102*(3), 757–771.
McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structural analyses. In R. Cudeck, S. H. C. du Toit, & D. Sörbom (Eds.), *Structural equation modeling: Present and future* (pp. 342–380). Lincolnwood, IL: Scientific Software International
McArdle, J. J., & Hamagami, F. (2001). Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 139–175). Washington, DC: American Psychological Association.
McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly*, *21*(4), 471–490.
McGeown, S., Goodwin, H., Henderson, N., & Wright, P. (2012). Gender differences in reading motivation: does sex or gender identity provide a better account? *Journal of Research in Reading*, *35*(3), 328–336.
McMillian, M. M., Frierson, H. T., & Campbell, F. A. (2011). Do gender differences exist in the academic identification of African American elementary school-aged children? *Journal of Black Psychology*, *37*(1), 78–98.

Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika*, *58*(4), 525–543.

Raag, T., Kusiak, K., Tumilty, M., Kelemen, A., Bernheimer, H., & Bond, J. (2011). Reconsidering SES and gender divides in literacy achievement: are the gaps across social class and gender necessary? *Educational Psychology*, *31*(6), 691–705.

Roberts, J. E., Burchinal, M., & Durham, M. (1999). Parents' report of vocabulary and grammatical development of African American preschoolers: child and environmental associations. *Child Development*, *70*(1), 92–106.

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (Dis) abilities: Evidence, theory and practice. In S. B. Neurman & D. K. Dickinson (Eds.), *Handbook of Early Literacy* (pp. 97–110). New York: Guilford Press.

Shaywitz, S. E., Shaywitz, B. A., Fletcher, J. M., & Escobar, M. D. (1990). Prevalence of reading disability in boys and girls: Results of the Connecticut Longitudinal Study. *JAMA*, *264*(8), 998–1002.

Tamis-LeMonda, C. S., Song, L., Leavell, A. S., Kahana-Kalman, R., & Yoshikawa, H. (2012). Ethnic differences in mother-infant language and gestural communications are associated with specific skills in infants. *Developmental Science*, *15*(3), 384–397. doi:10.1111/j.1467-7687.2012.01136.x

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech Language and Hearing Research*, *40*(6), 1245–1260.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–69.

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174.

Washington, J. A., Branum-Martin, L., Sun, C., & Lee-James, R. (2018). The impact of dialect density on the growth of language and reading in African American children. *Language Speech and Hearing Services in Schools*, *49*(2), 232–247. doi:10.1044/2018_LSHSS-17-0063