Revisiting the Fourth Grade Slump among Black Children: Taking a Closer Look at Oral Language and Reading

Nicole Patton Terry
Florida State University & Florida Center for Reading Research
https://orcid.org/0000-0002-4564-9394

Brandy Gatlin-Nash
University of California, Irvine
https://orcid.org/0000-0002-6925-8961

Mi-Young Webb
Coryus, LLC
https://orcid.org/ 0009-0006-4901-0674

S. Rebecca Summy
Florida State University & Florida Center for Reading Research
https://orcid.org/0000-0003-1809-8213

Rhonda Raines
Florida State University & Florida Center for Reading Research
https://orcid.org/0000-0003-0788-1501

Published online February 15, 2023

Contact Author:
Nicole Patton Terry, Ph.D.
Florida Center for Reading Research
2010 Levy Ave, Suite 100
Tallahassee, FL 32310

Acknowledgements:
This work was supported in part by funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health (1R24HDO75454-01), Institute of Education Sciences (R305B200020), and the Office of Special Education Programs (OSEP), U.S. Department of Education (H325D190037). The opinions expressed are ours and do not represent views of the funding agencies. We would like to thank the schools, teachers, children, and families who participated in this project, without whom this research would not have been possible. We would also like to thank our colleagues who provided feedback on this work, and The Dyslexia Foundation for the opportunity to share this important work.

Abstract

Nearly 30 years ago, Chall, Jacobs, and Baldwin (1990) introduced the *fourth-grade slump* to describe the unexpected deceleration of reading skills between first and fourth grades among children growing up in poverty and low-income households. Advances in our understanding of reading development and how race, racism, or other forms of discrimination are implicated in student achievement encourage another look at the slump among Black children. We reexamined the slump in a cross-sectional, heterogeneous sample (n = 757) of typically developing, Black children in grades 1- 4 (ages 6-11 years old), most of whom were growing up in low-income households. As in the original study, we observed a downward trend in word reading and reading comprehension skills between first and fourth grades. However, unlike the original study, children demonstrated relatively weaker performance on oral language measures at each grade level. Implications for studying oral language and reading in Black children are discussed.

Revisiting the Fourth-Grade Slump among Black Children: Taking a Closer Look at Oral Language and Reading

The *fourth-grade slump* is a term coined by Chall et al. (1990) to describe the sudden emergence of poor reading achievement among seemingly typically developing children who were growing up in poverty or low-income households who had demonstrated grade-appropriate reading skills earlier in elementary school. The seminal study spurred interest in closing the reading achievement gap between more and less affluent students. Many reasons that have been offered to explain the fourth-grade slump, including inefficient foundational reading skills, insufficient background and content knowledge, undetected reading disabilities, and lack of motivation to read (e.g., Guthrie & Humenick, 2004; Hirsch, 2003; Sanacore & Palumbo, 2009; Snow et al., 1998; Stockard, 2010). Cumulative and additive risk models have demonstrated that multiple factors can interact to either prevent and impede or promote and protect students from academic difficulty (Burchinal & Willoughby, 2013; Gutman et al., 2003). Therefore, it is plausible that these factors contribute to the slump, directly or indirectly. However, this study is focused on one factor that has received surprisingly little attention in discussions about the fourth-grade slump, despite growing empirical evidence supporting its substantial importance to reading development and achievement: oral language skills.

This oversight is particularly troublesome when considering approaches to supporting early reading and literacy skills among Black children, who present both researchers and educators with an interesting paradox. On the one hand, there is evidence to suggest that some oral language skills may be relative strengths among Black children that can be leveraged to support learning and achievement (e.g., Craig & Washington, 1994; Gardner-Neblett et al., 2017; Terry et al., 2018). One the other hand, national data continue to indicate that Black children are

more like to be growing up in historically marginalized communities where many people are living in poverty (e.g., Koball et al., 2021) and are more likely to be experiencing difficulty with reading in school (e.g., National Center for Education Statistics [NCES], 2019; 2022). For example, scores on the National Assessment of Educational Progress (NAEP) before the COVID-19 pandemic (NCES, 2019) revealed that only one in five fourth graders were eligible for free or reduced-price lunch programs and one in five Black fourth graders were reading at or above proficient levels. Trend data two years after the onset of the COVID-19 pandemic reveal similar patterns, as average performance declined more sharply for students in both groups (NCES, 2022).

The fourth-grade slump presents an interesting context within which to explore these seemingly complex and contradictory conditions for young Black learners. Therefore, in this descriptive study we revisited the fourth-grade slump within a large cross-sectional sample of Black children, many of whom were growing up in poverty and low-income households as indicated by their participation in federal free or reduced-price lunch programs. We focused on Black children for two reasons.

First, investigations of reading and literacy achievement *within* student subgroups can elucidate factors that promote and protect their healthy development, prevent difficulty from emerging, and characterize effective supports to recognize and respond to difficulty or disability when it arises (Cabrera & the SRCD Ethnic and Racial Issues Committee, 2013). Alternatively, focusing on between group comparisons can both mask the prevalence of reading difficulty among school-aged children in the United States and amplify the perception that children in certain subgroups are somehow deficient or deviant. Keep in mind: NAEP scores before the COVID-19 pandemic (NCES, 2019) revealed that, on average, only one in three White fourth

graders and one in two more affluent fourth graders were reading at or above proficient levels; their scores declined after the onset of the pandemic too (NCES, 2022). According to the NAEP, most fourth graders in US schools are not meeting grade level expectations for reading achievement in school, not just Black and Brown children and not just children growing up in poverty.

Second, though growing, there is relatively little empirical research examining within group variation in early oral language, reading, and literacy skills among students who are Black (Iruka et al., 2015, 2020; Washington et al., 2019). If we do not consider race as a social or structural construct that has theoretical or methodological implications for the study's design, results, or interpretation (e.g., Bruno & Iruka, 2022; Rad et al., 2018; Roberts et al., 2020), then there will be limitations on our shared understanding of how humans learn and develop and how sociocultural context shapes learning and development in both formal and informal interactions, including as children develop oral language, reading, and literacy skills in school.

Importantly, this study is not intended to be a replication of Chall et al (1990). Rather, guided by findings from the original study, our evolving understanding of how to position race in studies of child development, and accumulating research evidence on the relation between oral language, word reading, and reading comprehension skills over the last 30 years, this study is intended to expand and extend upon the field's understanding of the fourth-grade slump for Black children. To do so, we examined, first, whether a slump was evident, and second, if the slump could be linked to children's performance on measures of various aspects of oral language that contribute to word reading and reading comprehension.

Theoretical Foundations

This study is informed by ecological models of child development and developmental and cognitive models of reading. Ecological models, like Bronfenbrenner's bioecological model (e.g., Bronfenbrenner & Cici, 1994), posit that development progresses through interactions between individuals and their environment. Interactions, which happen over time and are influenced by the characteristics of the individual and the environment, are the primary mechanisms for development, with some contexts influencing more directly and others more indirectly. García Coll and colleagues have proposed an integrative model that expands ecological models to first consider socioculturally defined positionality factors (e.g., race, gender, social class) and their oppressive manifestation in environments (e.g., racism, discrimination, segregation) in children's development (e.g., García Coll, et al., 1996; Marks & García Coll, 2018). Doing so brings environmental factors that may be particularly meaningful for Black children to the forefront because it acknowledges that children's development is shaped by how children and families respond to oppressive conditions in their environment. Importantly, within an integrative model, child outcomes may be positive or negative, depending on interactions between factors within the individual and the environment. For reading and literacy achievement, these ecological models suggest that reading development can be impacted directly and indirectly by a number of factors at the individual, family, school, neighborhood, and broader community level, and that considering the impact of these factors, both individually and in concert as a system, can provide a better understanding of why some children struggle with learning to read (Keenan et al., 2006; Vukovic et al., 2013).

This study is also grounded in developmental and cognitive models of reading that account for factors that underlie reading directly. In particular, Chall et al. (1990) was grounded in Chall's stage model of reading development, where children advance from skills needed for

"learning to read" words to skills needed to understand text content to facilitate "reading to learn." Over time, both theory and research have continued to advance in ways that expand upon Chall's developmental model to understand cognitive and linguistic factors related to reading and literacy acquisition directly and indirectly (Semingson & Kerns, 2021). For example, the reading rope heuristic proposed by Scarborough (2001) reflects evidence from correlational and longitudinal research on concurrent and predictive associations between various early oral language and word recognition skills and later word reading and text comprehension in the primary grades, including background knowledge, vocabulary, syntax, phonological awareness, and decoding. More recently, the direct and indirect effects model of reading (Kim, 2020) specifies the hierarchical, interactive and dynamic processes, skills, and structural relations involved in successful reading comprehension, including word recognition, listening comprehension, reading fluency, background knowledge, socioemotions, higher-order cognitions, vocabulary, syntactic knowledge, phonology, morphology, orthography, and executive function.

It is important to note that there is cross-cultural and cross-linguistic evidence in support of many of these models (see Cain et al., 2017 for a more thorough discussion on theories of reading). In their focus on cognitive, linguistic, and psychological factors that underlie basic reading processes and development independent of environmental contexts, these models remain critical to our understanding of how reading develops typically and atypically. However, much of the research upon which these models were developed or extended was not conducted with children from historically and systemically marginalized groups or in historically and systemically marginalized schools and communities, including Black children and children growing up in poverty (Artiles et al., 1997; Lindo, 2006; Reed et al., 2012). Thus, our

understanding of their application to addressing reading difficulty among Black children is limited.

Overview of the Original Fourth Grade Slump Study

The original study from which the fourth-grade slump emerged was developed and carried out in response to requests from the National Institute of Education for investigations focused on the education of children growing up in what were described as "low socioeconomic status (SES) conditions", with a particular focus on conditions in schools and in homes (Chall & Snow, 1982). The sample included 30 children—10 in second grade, 12 in fourth grade, and 8 in sixth grade—who were tested and retested one year later in third, fifth, and seventh grade, respectively, with performance compared with test norms. The investigators were particularly focused on why literacy achievement gaps between more and less affluent students seemed to grow as children progressed through school.

The study was carefully planned to control some variables, including family income and student reading ability. For example, the research team recognized variation in the category of "low SES household" and took steps to capture this variability. As such, the families were diverse and included single-parent and two-parent families, multilingual and multigenerational families, and families with parents who attended college and parents who did not graduate from high school. The research team also attempted to control for reading ability. The study included only children whose reading performance in school varied within the average range to better uncover home and school factors that contributed to reading performance. The study did not include children whose reading performance was very high or very low due to disabilities or other conditions that may have been related to their reading performance. The research team did not control for race. The sample was racially diverse and included children living in three

low SES neighborhoods in a large metropolitan area—two that were racially and ethnically diverse and one that was predominantly White. Schools attended by the participants also varied in their racial composition, including schools that were predominantly Black and others that were predominantly White.

Analyses of student performance in the cross-sequential longitudinal design produced several key findings. As reported most frequently, the slump seemed to occur as children transitioned from "learning to read" to "reading to learn", or stages 2 and 3 in Chall's developmental model of reading (Chall, 1983). Children's performance on various reading measures was commensurate with test norms in grades 2 and 3, began decelerating in grade 4, and continued to fall through grade 7. The slump was first observed and most pronounced on a measure of vocabulary, where children were asked to define words of increasing difficulty. By grade 4, children were performing 1 year below test norms; by grade 7, they were performing two or more years below. This slump in vocabulary performance was followed by a slump in word reading in grade 4, spelling in grade 6, and reading comprehension in grade 7. Importantly, there was a disconnect between children's grades and performance on specific study measures: approximately 75% of children had As and Bs on their report cards (including children who were already performing below average on study measures) and approximately 68% of mothers reported that their children were reading above grade-level expectations.

Though robust in its longitudinal approach, the small sample size limited the generalizability of the findings. Nonetheless, the study's findings have continued to inform reading and literacy research, practice, and policy. Importantly, the observed slump in reading and literacy achievement did not indicate a loss of skills, but rather a reduction in the rate of growth that is commensurate with grade-level expectations. This finding suggests that, if

addressed effectively, the slump could be prevented, stopped, or reversed. Chall et al. offered one primary recommendation based on their findings: that children growing up in low-SES conditions be provided with the same literacy instruction as students growing up in more affluent conditions.

The Importance of Considering Race: Intersections with Poverty, Reading, and Literacy

One strength that is clear from the study's design and the interpretation of the findings is the research team's attempts to take a strengths-based approach to investigating reading and literacy among children growing up in poverty. Rather than simply reporting on income-based achievement gaps, the researchers sought to capture the child, school, and family factors that mattered for reading and literacy performance. In fact, one stated goal of the study was to better understand how some children from low SES homes succeeded in school and what their families did to enable their children's success. The researchers' insistence that children growing up in poverty were not different from children growing up in middle SES and more affluent households was controversial at the time.

Nonetheless, one limitation of the study, specifically for Black children, is its color-blind approach (Bruno & Iruka, 2022). By not capturing and operationalizing race in the study, the social, cultural, and historical impact of race on the children's developmental outcomes was not documented well and not considered fully in the design of the study or the interpretation of the findings. The sample size would have prevented such analyses. However, in the decades since the original study, researchers have observed differences in various factors related directly and indirectly to reading, oral language, and literacy outcomes for Black children. For example, several literacy-focused ethnographic studies (e.g., Ladson-Billings, 2009) and quasi-experimental studies (e.g., Cartledge et al., 2015) have reported on culturally responsive and

affirming interactions between Black teachers and school leaders and Black students and their families that supported student performance in school. Larger quasi-experimental and experimental studies have revealed important race and race-related effects for Black children's school achievement and success, including that Black students who have at least one Black teacher in kindergarten through grade 3 were significantly more likely to graduate from high school and to enroll in college (Gershenson et al., 2018) and that teachers vary significantly in their race-based implicit and explicit biases about students and that this variation is associated significantly with racial disparities in teacher expectations for student attainment, student test scores, and student suspension rates (Chin et al., 2020; Gershenson et al., 2016). Certainly, studies like these were well beyond the scope of the original fourth grade slump study. Nevertheless, these more recent findings suggest that much more could be learned by examining the intersectionality of poverty with several factors within the sample, including race, racism, and other forms of discrimination encountered by children and families in the northeastern United States in the early 1980s.

Moreover, race and poverty were and continue to be confounded in the United States (Koball et al., 2021). Black children continue to be much more likely to be growing up in poverty and low-income households, and these income differences are salient in their homes, schools, and communities. "Ordinary" childhood and school experiences for Black children are affected by race. Scholars have argued that, because there is no biological basis for race, it is more appropriate not to consider race in scientific research; however, doing so ignores the realities of how race has been used historically in the United States as a social, cultural, and political tool to create systems and structures that result in social, economic, health, educational, and other disparities between racial groups (e.g., DeCuir-Gunby & Schutz, 2014; Shi & Stevens,

2021; Smedley & Smedley, 2005; Terry et al., 2022). Albeit aspirational, policy and practice recommendations like those from the original study for improved reading and literacy experiences for *all* children that do not address racial and economic disparities (and how they are enacted through racism and other forms of discrimination) will be limited in their successful implementation in classrooms, schools, and communities with Black children and children growing up in poverty.

Revisiting Language and Its Importance to the Slump among Black Children

There is good reason to reexamine the contribution of oral language skills to the fourth-grade slump, especially among Black children. First, as noted previously, Black children are disproportionately more likely to experience difficulty with reading in the primary grades (NCES, 2019; 2022) and to be represented among families living in poverty and in low-income households (Koball et al., 2021). Researchers have observed that the quantity and quality of language children experience in homes and in schools is associated with poverty and that these poverty-influenced differences are associated with and predictive of reading achievement in school (e.g., Golinkoff et al., 2018; Logan et al., 2019). That said, there is heterogeneity among Black children—they are not all alike. Within-group examinations of language and reading among school-aged Black children are rare, rendering this student population understudied and limiting the availability of solutions to address their specific needs.

Second, academic vocabulary (words often encountered in text but not generally used in everyday conversation) was one of two primary causes Chall et al. (1990) offered in explaining the slump. Chall and colleagues included other language measures in the study. However, in contrast to observations of growing difficulty with less common words, the researchers found that student performance on these measures remained commensurate with test norms, with no

significant differences between students performing above or below average in reading. Language difficulties were only noted for measures of academic vocabulary. There is consensus within the field that variation in vocabulary is associated with reading achievement and that many children do not have the vocabulary knowledge necessary to support advanced word reading and reading comprehension (e.g., Castles et al., 2018).

However, there is also growing research evidence that variability in language ability, above and beyond vocabulary, is associated with advanced word reading and reading comprehension skills (Foorman et al., 2015; NELP, 2009; NICHD Early Child Care Research Network, 2005; Pearson et al., 2020). For example, recent findings from the Language and Reading Research Consortium (LARRC) indicate that early language development from preschool through the primary grades is dimensional, with vocabulary, grammar, and higher-level language skills (i.e., comprehension monitoring, narrative structure, and inferencing) contributing uniquely to reading comprehension in third grade (LARRC, 2015, 2017) and lexical quality (i.e., knowledge of a word's vocabulary, orthography, morphosyntax, and phonology) in preschool predicting reading in first grade (Murphy et al., 2016). Similar findings have been reported by Cabell et al. (2011, 2013), who found distinct emergent literacy profiles among preschoolers from low-income households; children with lower oral language skills demonstrated weaker early reading performance on code-related tasks. A common finding across these studies is that variation in oral language skills matters for children who struggle with reading.

Finally, specifically with regard to Black children, researchers have also observed strengths in oral language skills within this student population, in particular related to narrative, decontextualized language, and complex syntax (e.g., Craig & Washington, 1994; Gardner-

Neblett & Iruka, 2015; Mills, 2015). However, there are simply few studies with large, representative samples of Black children that examine relations between performance on various reading and language measures. One exception is studies that have focused on one aspect of Black children's language that has received considerable attention in the literature: spoken dialect variation. Many Black children come to the task of learning to read speaking African American English (AAE), a nonmainstream American English (NMAE) dialect whose language form, content, and use varies from Mainstream American English and, importantly, from standard English orthography (Terry et al., 2018). Accumulating empirical evidence has revealed significant concurrent relations between the density of children's spoken NMAE use and their performance on reading and literacy tasks (e.g., Gatlin & Wanzek, 2015; Terry et al., 2012). However, causal relations are inconsistent. with some finding reciprocal relations between reading and spoken dialect use (Washington et al., 2018), others finding more complex U-shaped relations with dense speakers demonstrating stronger reading performance than less dense speakers (e.g., Connor & Craig, 2006), and others finding stronger relations between change in spoken dialect use (i.e., dialect-shifting) and reading than their density of use at one time point (e.g., Craig et al., 2009). Thus, although denser AAE speakers tend to demonstrate weaker performance on language and reading measures, there is not sufficient evidence to suggest that speaking AAE causes difficulty in language or reading achievement (Terry et al., 2018).

Purpose of the Study

In response to calls for research on the language abilities of diverse children that is both strengths-based and race-conscious, researchers have begun to build a more nuanced and representative evidence base for language ability in young children (e.g., Logan et al., 2019; Purpura, 2019; Washington et al., 2018; 2019). Specifically for Black children, a more

comprehensive understanding of language could improve our understanding of and response to reading difficulties observed in this student population. More specifically for this study, it is plausible that the reading difficulty captured by the fourth grade slump could be associated with and alleviated by addressing differences in oral language ability more broadly.

That said, there are important similarities and differences between this study and the original fourth grade slump study. Like Chall et al. (1990), this study included only children who were growing up in poverty and low-income households and who were not receiving additional services and supports in school for reading or language. We also administered multiple norm-referenced, standardized oral language and reading measures and compared student performance with test norms. Unlike Chall et al. (1990), the sample in this study was large (with nearly 150 or more children at each grade level), included only Black children, and included data at only one point in time (the data were not longitudinal). Thus, the sample in this study may be more representative of Black children in the primary grades in public schools in the United States.

In sum, the aim of this brief descriptive study was to determine whether trends seen in Chall et al. (1990) would emerge in a heterogenous sample of typically developing Black children who were growing up in poverty or low-income households: Did children develop early literacy and language skills at a normative pace until third grade? Did their language and literacy achievement begin to slump in fourth grade? If so, was oral language ability associated with performance, and were there trends in the data that point to skills that might be appropriate targets for support to alleviate downward trends in reading achievement?

Methods

Participants

The study sample was drawn from a larger sample of Black students enrolled in first through fourth grades in a large urban school district in the southeastern United States. Students were enrolled in one of five public schools where approximately 87% -100% of children participated in free or reduced-priced lunch programs. The schools were located in predominantly Black neighborhoods where publicly available census data during the time of the study indicated that, during the previous 12 months, the majority of children under 18 years old in the area were living in households where family income was at or near the federal poverty threshold (approximately 60% across neighborhood zip codes), the majority of adults above 25 years old had achieved less than a high school degree or GED (approximately 56% across neighborhood zip codes), and the majority of residents above 16 years old were unemployed (approximately 60% across neighborhood zip codes) (US Census Bureau, 2015).

This study included students whose performance on the Kaufman Brief Intelligence Test, Second Edition (KBIT- 2; Kaufman & Kaufman, 2004) was within the average range (M = 97.46, SD = 13.77) and who were not receiving special education services. In addition, for descriptive purposes, spoken dialect variation was measured using the Diagnostic Evaluation of Language Variation Screening Test (DELV-ST; Seymour et al., 2003). As has been done in previous studies (Terry et al., 2012), we used children's responses to Part 1 to compute a percentage dialect variation score (DVAR), which represents rate of NMAE production.

The final study sample included 757 students (398 girls) whose ages ranged from 6.11 years to 11.89 years old (M = 8.19 years, SD = 1.22 years). Table 1 presents demographic information of the participating students for each grade. The grades differed marginally in performance on the KBIT-2, (F (3) = 2.61, p = .05). In addition, average rates of NMAE production were highest in grade 1, decreasing by 10% in grade 2, and remained the same in

grades 3 and 4.

Measures

Oral Language. Three subtests from each of the Test of Language Development Primary-Fourth Edition (TOLD-P4; Hammill & Newcomer, 2008a) and the Test of Language Development Intermediate-Fourth Edition (TOLD-I4; Hammill & Newcomer, 2008b) were used to assess students' oral language skills. The TOLD-P4 is designed for children younger than 8 years of age, while TOLD-I4 is designed for children who are 8 years and older. Hence, the majority of participants in grades 1 and 2 were administered the morphological completion, syntactic understanding, and picture vocabulary subtests of the TOLD-P4, and the majority of participants in grades 3 and 4 were assessed using the morphological comprehension, sentence combining, and picture vocabulary subtests of the TOLD-I4.

The TOLD-P4 picture vocabulary subtest is a receptive measure of children's understanding of spoken words. Children are presented with an array of four pictures and asked to identify the one that depicts a word spoken by the assessor (e.g., "See these pictures? Show me baby, point to baby."). The TOLD-P4 syntactic understanding subtest is a receptive measure of children's ability to comprehend the meaning of sentences. Children are asked to select one of three pictures that most accurately represents the stimulus sentence (e.g., "Point to the picture that matches 'The ball is round.' "). The TOLD-P4 morphological completion subtest is an expressive measure of children's ability to recognize, understand, and use common English morphological forms in a cloze format. Children are presented with a picture and asked to provide a single word to complete a sentence (e.g., "Bill is a boy and John is a boy. They are both__."). Internal consistency reliability (measured by Cronbach's alpha) for this sample was .78 for picture vocabulary, .67 for syntactic understanding, and .91 for

morphological completion.

The TOLD-I4 picture vocabulary subtest is a receptive measure of children's ability to understand two-word phrases. Given a set of six pictures, children are asked to identify (by pointing) the picture that represents the two-word stimulus (e.g., "Point to the picture that best goes with these words, 'monkey see.' "). The TOLD-I4 sentence combining subtest is an expressive measure of children's ability to use phrases, embedded clauses, transformations, and adjectives to construct sentences. Children are asked to create a compound sentence from two or more simple sentences presented verbally (e.g., Make one sentence out of the sentences. Make your sentences as short as you can. 'I am big. I am tall.' "). The child is required to include all key elements in the stimulus sentences (i.e., "I am big and tall."). The TOLD-I4 morphological comprehension subtest is a receptive measure of grammatical knowledge. Children are asked to identify grammatically correct and incorrect sentences presented orally (e.g., "You tell me if the sentence I say is correct or incorrect. 'I see you.' Is that sentence correct or incorrect?"). Incorrect sentences contain morphological errors such as in the noun-verb agreement, pronouns, negatives, or plurals. On each subtest, raw scores are transformed into scaled scores (M = 10, SD = 3). Internal consistency reliabilities for our sample were .91, .89, and .92 for sentence combining, picture vocabulary, and morphological comprehension, respectively.

Phonological Awareness. The elision and blending words subtests of the Comprehensive Test of Phonological Processing, Second Edition (CTOPP; Wagner et al., 2013) were used to measure students' phonological awareness. On the elision subtest, children were asked to say a word and then say what is left after removing sounds (e.g., "Say farm. Now say farm without saying /f/."). On the blending words subtest, children were asked to

combine spoken sounds to form words (e.g., "What word do these sounds make? s – ŭn"). On each subtest, raw scores are transformed into scaled scores (M = 10, SD = 3). Internal consistency reliability for the sample was .84 for Blending Words and .93 for Elision.

Reading Skill. Reading ability was measured with four subtests of the Woodcock-Johnson Diagnostic Reading Battery, Third Edition (WJ-III; Woodcock et al., 2004). On the letter-word identification subtest, children's letter and word identification skills were assessed by asking them to identify letters and single words of increasing difficulty. On the word attack subtest, children's phonic and structural analysis skills were assessed by asking them to read unfamiliar nonwords. On the passage comprehension subtest, children's understanding of written text was assessed by asking them to match phrases to the appropriate picture in an array and provide a missing word to complete sentences and brief paragraphs. Finally, on the reading fluency subtest, children's fluency was assessed by asking them to read simple sentences and decide whether the statement was true or false within 3 minutes. On each subtest, raw scores are transformed into standard scores (M = 100, SD = 15). Internal consistency reliability for subtests within the current sample was .95 for letter word identification, .94 for word attack, .89 for passage comprehension, and .96 for reading fluency.

Students were also given the Gray Oral Reading Tests, Fifth Edition (GORT-5; Wiederholt & Bryant, 2012), a measure of reading accuracy, fluency, and comprehension. On this task, children are asked to read the passage orally as carefully and quickly as they can while the assessor times the child's reading and records errors. After reading the passage, the assessor removes the passage and asks comprehension questions. Performance is scored for rate, accuracy, fluency, and comprehension (scaled scores, M = 10, SD = 3). In this study, we

used the normative composite score, Oral Reading Index score (M = 100, SD = 15, *Average* = 90-110), which is calculated from the sum of the fluency and comprehension scaled scores. We used this score both because it does not have floor or ceiling effects at any age level and because it represents a "capstone" reading measure. That is, students who achieved average Oral Reading Index scores could be characterized as having integrated word reading, fluency, and comprehension skills that were applied efficiently to read and understand the text. Internal consistency reliability for the sample was .79.

Procedures

This study was approved by the university's institutional review board for the appropriate conduct of research with human subjects. Each assessment was administered individually in a quiet setting in the child's school. Trained graduate research assistants administered and scored the assessments according to standard administration procedures specified in the manual. Scoring and data entry reliability were established for approximately 20% of the data for each assessment. Raw scores were automatically calculated by each item score entry, and standardized scores were also automatically calculated via SAS programming, except for WJ-III for which we used the test's scoring software. Specifically, SAS programming was written to look up the scores that matched a participant's chronological age and the raw scores in scoring tables that were created based on the assessments' manuals. Agreement for entry reliability was 92% on average, and scoring reliability achieved 100% using autoscoring system and double scoring for WJ-III.

Results

Table 2 displays average student performance on each of the measures by grade level, as well as the results from independent t-tests comparing average performance in grades 1 and 4.

Ranges in performance and the percentage of students performing one or more standard deviations below the mean were also provided to elucidate the heterogeneity of the sample. Indeed, at each grade level, there were students who were performing well above and below age expected norms on each reading and language measure. In addition, the percentage of students performing below grade level on the reading measures varied as expected, with as few as 4% of students in grade 2 not meeting expectation on the WJ-III word attack subtest and as many as 26% of students in grade 4 not meeting expectation on the WJ-III passage comprehension subtest. More variability was observed on the language measures, with as few as 18% of students in grade 1 not meeting expectation on the TOLD syntax subtest and as many as 80% of students in grade 3 not meeting expectation on the CTOPP blending words subtest. In grades 3 and 4, at least 40% or more students in the sample were performing one standard deviation or more below the mean on every language measure.

Results from a series of independent t-tests revealed significant differences in average standard or scaled score performance for students in grades 1 compared to grade 4. All tests were significant at the p-value less than .001 (see Table 2). Mean scores indicated that, on average, students in grade 1 had high average reading skills, average oral receptive vocabulary and oral syntax skills, and low average oral morphology and phonological awareness skills. Students in grade 4 had low average performance on all reading measures and below average performance on all oral language measures—their scores were one standard deviation below the test mean. All in all, average standard scores on the reading measures decreased between grades 1 and 4. Although within the average range, the sample mean on reading measures declined by 11.8 to 13.7 standard score points between grades 1 and 4. Scores on the WJ-III passage comprehension subtest showed steady declines across the grades, with a difference of approximately four

standard score points for each grade level. Similarly, average scaled scores on the oral language measures also decreased between grades 1 and 4. Scaled scores on the CTOPP blending words subtest declined nearly one point between grades 1 and 2 and between grades 3 and 4, and the average scaled scores in grades 2 and 4 remained one standard deviation below the normative test sample means. Scaled scores on the CTOPP elision subtest showed steady decline across the grades at a rate of approximately 0.4 to 0.5 standard score points each grade level. Finally, a similar pattern was observed for the capstone reading measure. On average, students in each grade level had GORT Oral Reading Index scores that were below the test norm average.

Table 3 presents correlations among the scores on the reading and oral language measures across all grades, adjusted for age. Correlations among the reading measures were relatively strong (ranging from .70 to.87), whereas correlations among the oral language measures were moderate (ranging from .46 to .58). Correlations between scores on the reading and oral language measures ranged from .40 (CTOPP blending words and GORT-5 comprehension; morphology and reading fluency) to .68 (elision and letter-word identification).

Figures 1-3 provide a visual representation of these relations. In each instance, the relation that is depicted is that between a reading skill and an oral language skill that supports it directly. Accordingly, what is shown are associations between the CTOPP elision and blending subtests (phonological awareness) and the WJ-III letter-word identification (word reading, Figure 1), Word Attack (decoding, Fig. 1), and reading fluency (reading accuracy, Fig. 2) subtests, and between the TOLD vocabulary, morphology, and syntax subtests and the WJ-III passage comprehension (reading comprehension, Fig. 3) and GORT-5 total score (reading comprehension, Fig. 3). The bubble represents the sample's average scaled score on the measure on the x-axis at each grade level. The location of each bubble was determined by the sample's

average score on the measure on the y-axis at each grade level. The size of the bubble represents the number of children in each grade relative to the total number of children in the data; the larger the bubble, the more children in that grade level who fall into that position. The number next to each bubble is the grade level.

The figures clearly show a decline in average reading and oral language scores, as the bubbles fall to and below the regression line and into the lower right quadrant from grade 1 (usually the highest bubble on the figure) to grade 4 (usually the lowest bubble on the figure). Take for example, Figure 3. The relation between scores on the WJ-III passage comprehension and TOLD morphology, syntax, and vocabulary (top row) largely deviates from the perfect regression line, indicating that reading comprehension declined more dramatically between grades 1 and 4 relative to the decline in its related oral language skills. In general, positions on the bubble plot begin to deviate from the perfect regression line in grade 3 and exist prominently below the line in grade 4. An even more dramatic pattern was found for the GORT-5 (Fig. 3, bottom row), where all bubbles for each grade are well below the regression line and located in the left bottom quadrant. A similar pattern was also found for the word reading skills, where deviations from the regression line were most noticeable in grade 4 for WJ-III letter word identification (Fig. 1, top row) and word attack (Fig. 1, bottom row) and in grades 3 and 4 for reading fluency (Fig. 2). Interestingly, children demonstrated relatively strong word-level skills but weak phonological awareness skills. Hence, unlike reading comprehension in Figure 3, the bubbles in Figure 1 are in the upper-left quadrant. Finally, it is quite startling that, in each figure, the bubbles are always in the upper- or lower-left quadrants, indicating that, on average, students' oral language performance did not meet the test average at any grade level.

Discussion

Framed within the context of the classic fourth grade slump study (Chall et al., 1990), the primary purpose of this descriptive study was to examine oral language and reading skills among Black children in grades 1 through 4. Like the original study, the sample included only typically developing children who were growing up in poverty and low-income households and focused on within group variation by comparing student performance to test norms. Unlike the original study, the sample included a relatively large number of Black students measured at one point in time. Though not longitudinal, this study sought to capitalize on advances in our understanding of the relation between oral language and reading while also focusing on within group variation among Black students at each grade level. Three primary findings emerged that both confirm and expand those reported in the original study.

First, the data reveal much heterogeneity among Black children in the primary grades. Some students were excelling in oral language and reading skills at each grade level—an important finding given stereotypical reports of widespread underachievement among Black children who are growing up in poverty and low-income households and who are enrolled in schools where the majority of students are Black and growing up in marginalized, underserved communities. At each grade level, the mean score on reading tests were within or above the average range. The data also reveal that many students were experiencing difficulty with specific reading and oral language skills. On each measure, there were students who were performing one or more standard deviations below the mean at each grade level—an important finding given reports that many Black students are disproportionately less likely to be identified for disabilities and placed in special education (Morgan et al., 2017). We would be remiss if we did not point out that, based on their test scores alone, some students in our study may meet criteria for eligibility to receive special education services in school. Yet, none were. The misrepresentation

of Black students in special education is a long-standing, complex, and multidimensional problem of educational equity (Skiba et al., 2016). It is important to keep in mind that eligibility for special education services is determined based upon multiple data sources. It would be inappropriate to interpret these test scores as an indication that a large percentage of students in the sample had undiagnosed reading disabilities or specific oral language impairments. However, the data do suggest that a closer look is warranted.

Second, similar to Chall et al. (1990), the data suggest a downward trend in word reading and reading comprehension skills between grades 1 and 4. On average, students in grade 1 performed well within the average range on multiple reading measures, including decoding, word reading, reading comprehension, and reading fluency. In grade 4, average performances were 12 to 14 standard score points below those of children in grade 1. Differences were most pronounced for word-level skills between grades 2 and 3 (e.g., decoding and single word reading) and for text level skills between grades 3 and 4 (e.g., fluency and comprehension). Keeping in mind that these reading tasks become increasingly difficult as the words and text become less familiar, more complex, and more varied, these results suggest that by grade 3, some students in this sample may not have had the word reading skills necessary to read extended texts with a level of accuracy and automaticity that supports comprehension. Although word recognition skills remained relatively strong, the downward trend continued to be evident in grade 4, as the demand for comprehending texts increased. Thus, as observed in the original study, the deceleration in word reading skills in the earlier grades was followed precipitously by a sharp decline in reading comprehension in the later grades. It is important to note that in grade 4, many children were performing in the average range on these reading measures. However, their performance on the GORT-5

suggests that these skills were not well-integrated and not functioning in a manner that supported grade-appropriate reading ability. Keeping in mind that average performance is 90-110, average student performance on the GORT Oral Reading Index was at or below the test norm at every grade level.

Third, unlike Chall et al. (1990), students in grade 1 did not demonstrate performance on oral language measures that was consistent with test norms. On average, students in first grade demonstrated low average phonological and morphological skills and average vocabulary and syntax skills. Phonological skills were no longer in the average range for students in grade 2. In addition to below average phonological skills, students in grade 3 also demonstrated below average morphology and syntax skills. In grade 4, average performance on all oral language measures was one standard deviation below the test mean. This pattern is different from that reported by Chall et al. (1990), who observed children's performance on oral language measures to be commensurate with test norms, except for a deceleration in vocabulary scores that was first observed in grade 4. In contrast, the results of this study suggest that relatively stronger oral vocabulary skills were not sufficient enough to promote expected word reading, comprehension, or fluency skills at any grade level. In fact, we observed little to no change in scores on oral language measures, on average. Thus, children in this sample may be better characterized not as exhibiting a slump or deceleration but rather stagnant oral language development. As shown in Figures 1-3, minimal change in children's oral language skills was associated with lower average reading performance at each grade level.

Although different, these results align well with current models of reading, which reflect complex contributions of oral language to reading that are above and beyond

vocabulary and include grammar, inference, and discourse skills to gain meaning from texts (LARRC, 2015; 2017). For example, in a large longitudinal sample of children who were followed from preschool to grade 3, the NICHD Early Child Care Research Network (2005) found that a comprehensive measure of oral language (i.e., syntax, morphology, vocabulary, and communication abilities) predicted word reading in grade 1 and reading comprehension in grade 3, above and beyond a measure of vocabulary alone. Arguing for a more comprehensive view of oral language and its relation to reading, the researchers noted that instruction focused only on vocabulary would be limited and not result in strong reading achievement. In this study, oral vocabulary was the strongest oral language skill for most children, perhaps reflecting a more intentional focus on vocabulary instruction in schools than other oral language skills (Hirsch, 2003). Thus, in this sample, it may be that relatively weaker reading performance in grade 4 could be explained, in part, by insufficient growth in oral language skills to support more advanced reading skills.

Conclusion and Implications

In this study, we revisited the fourth-grade slump with a renewed lens, focusing on within group variation among Black children growing up in poverty and low-income households. Guided by the original study (Chall et al., 1990) and informed by current advances in both reading and oral language research, we questioned whether the slump would still be observed, and if so, how it might manifest within this understudied student population. In some ways, the results are neither novel nor unexpected: average performance on norm-referenced standardized measures of specific reading skills was higher among students in grade 1 than in grade 4. However, the downward trend in average reading performance across grade levels appeared to be associated not with a lack of foundational reading or oral

vocabulary skills, but with less well-developed oral language skills.

This study is not without limitations. Obviously, replication of the findings within a longitudinal sample would provide more stringent support for what appears to be a slump in reading skills. The data reported here are means for students separated by grade level at one point in time; a true slump can only be observed over time. Moreover, although these data suggest there may be individual differences in both oral language and reading skills within this student group, a person-centered approach to the data would be necessary to determine if development of these skills proceeds differently in specific subgroups (Cabell et al., 2011, 2013). It is plausible there are systematic individual differences among Black children growing up in poverty, which can be illuminated with large representative study samples that capture heterogeneity. This study sample was large, but that does not mean it is representative of all Black children.

In addition, an examination of the role of spoken dialect variation in student performance on the expressive oral language measures used in this study is warranted. Protocols were in place to minimize scoring bias, according to directions within the test manuals. Moreover, average student performance was similar on both expressive and receptive oral language tasks. Nonetheless, it is plausible that student performance reflected scoring bias on items that are sensitive to spoken dialect differences (Hendricks & Adlof, 2017). Finally, data were not available to capture conditions in classrooms, schools, homes, and communities that may have contributed to student performance on the oral language and reading measures. Although less often reported, there were findings from the original fourth grade slump study focused on conditions in homes and schools that were associated with stronger reading and literacy performance (Chall & Snow, 1982; Snow et al., 2013). More contemporary studies

have also explored these relations (e.g., Chin et al., 2020; Gatlin-Nash et al., 2021; Gershenson et al., 2016, 2018), and additional research would be beneficial.

These limitations notwithstanding, these findings add to the literature on reading development and achievement for Black children and point to directions for research and practice. All in all, the results support a comprehensive approach to oral language development and instruction that supports more advanced word reading and comprehension skills. They also point to a more holistic appreciation for how children vary. Depictions of "at-risk" samples of Black children and children growing up in poverty are too narrow to capture development in a manner that leverages their strengths and challenges with strategic intention. These student populations are not homogeneous and should not be treated as such in research or in practice.

Finally, the results illuminate what can be learned from investigations that focus on understudied groups. We think it is worth asking, would we learn anything new if Chall et al. (1990) conducted their study today, armed with all we have learned about oral language, reading, race, and poverty over the last 30 years and under the current conditions in which we expect children to learn to read and succeed in school? We imagine that we would come to the same conclusion: the provision of high-quality, effective reading and literacy instruction should be standard practice for all learners. However, we also hope we would agree that ignoring the very tangible forces that racism and poverty impart upon children's lives would limit our research, its interpretation, and its ability to inform policies and practices that can promote and protect reading achievement for students from historically marginalized groups, including Black children. And we would fight for them, together.

References

Artiles, A. J., Trent, S. C., & Kuan, L.-A. (1997). Learning disabilities empirical research on ethnic minority students: An analysis of 22 years of studies published in selected refereed journals. *Learning Disabilities Research and Practice, 12*(2), 82–91.

Bruno, E. P., & Iruka, I. U. (2022). Reexamining the Carolina Abecedarian project using an antiracist perspective: Implications for early care and education research. *Early Childhood Research Quarterly*, *58*, 165-176.

Cabell, S.Q., Justice, L.M., Konold, T.R., and McGinty, A.S. (2011). Profiles of emergent literacy skills among preschool children who are at risk for academic difficulties. *Early Childhood Research Quarterly, 26 (1),* 1-14.

Cabell, S. Q., Justice, L. M., Logan, J. A. R., & Konold, T. R. (2013). Emergent literacy profiles among prekindergarten children from low-SES backgrounds: Longitudinal considerations. *Early Childhood Research Quarterly, 28*, 608–620.

Cabrera, N. J., & The SRCD Ethnic and Racial Issues Committee. (2013). Positive development of minority children. *Social Policy Report, 27(2*), 1–22.

Cartledge, G., Keesey, S., Bennett, J., Gallant, D., & Ramnath, R. (2015). Effects of culturally relevant materials on the reading performance of second-grade African Americans with reading/special education risk. *Multiple Voices, 15(*1), 22-43.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest, 19(1*), 5-51.

Chall, J. S. (1983). *Stages of reading development.* New York: McGraw-Hill.

Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind.* Cambridge, MA: Harvard University Press.

Chall, J., & Snow, C. (1982). *Families and literacy: The contribution of out-of-school experiences to children's acquisition of literacy*. Final Report.

Chin, M. J., Quinn, D. M., Dhaliwal, T. K., & Lovison, V. S. (2020). Bias in the air: A nationwide exploration of teachers' implicit racial attitudes, aggregate bias, and student outcomes. *Educational Researcher*, 49 (8), 566-578.

Cain, K., Compton, D. L., & Parrila, R. K. (2017). *Theories of Reading Development.* John Benjamins.

Connor, C., & Craig, H. (2006). African American preschoolers' language, emergent literacy skills, and use of AAE: A complex relation. *Journal of Speech, Language, and Hearing Research*, *49,* 771-792

Craig, H. K., & Washington, J. A. (1994). The complex syntax skills of poor, urban, African-American preschoolers at school entry. *Language, Speech, and Hearing Services in Schools*, *25*(3), 181-190.

Craig, H. K., Zhang, L., Hensel, S. L., & Quinn, E. J., (2009). African American English-speaking students: An examination of the relationship between dialect shifting and reading outcomes. *Journal of Speech, Language and Hearing Research*, 52, 839-855.

DeCuir-Gunby, J. T., & Schutz, P. A. (2014). Researching race within educational psychology contexts. *Educational Psychologist, 49* (4), 244–260.

García Coll, C., Lamberty, G., Jenkins, R., McAdoo, H., Crnic, K., Wasik, B., & García, H. (1996). An integrative model for the study of developmental competencies in minority children. *Child Development, 67*, 1891–1914.

Gardner-Neblett, N., Curenton, S. M., & Blitch, K. A. (2017). Viewing African American

children's oral language skills as a strength. In *African American children in early childhood education*. Emerald.

Gardner-Neblett, N., & Iruka, I. U. (2015). Oral narrative skills: Explaining the language-emergent literacy link by race/ethnicity and SES. *Developmental Psychology*, *51*(7), 889.

Gatlin, B., & Wanzek, J. (2015). Relations among children's use of dialect and literacy skills: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 58*(4), 1306-1318.

Gatlin-Nash, B., Hwang, J.K., Tani, N.E., Zargar, E., Wood, T.S., Yang, D., Powell, K.B., Connor, C.M. (2021). Using assessment to improve the precision of teachers' perception of students' academic competence. *The Elementary School Journal*, 121 (4), 000.

Gershenson, S , Hart, C. M. D. , Hyman, J. , Lindsay, C. , & Papageorge, N. W. (2018). *The long-run impacts of same-race teachers*. National Bureau of Economic Research.

Gershenson, S., Holt, S. B., & Papageorge, N. W. (2016). Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Economics of Education Review*, *52*, 209-224.

Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2018). Language matters: Denying the existence of the 30-million word gap has serious consequences. *Child Development, 90(3),* 985-992.

Guthrie, J.T., & Humenick, N.M. (2004). Motivating students to read: Evidence for classroom practices that increase reading motivation and achievement. In P. McCardle & V. Chhabra (Eds.), The voice of evidence in reading research (pp. 329-354). Brookes.

Hammill, D. D., & Newcomer, P. L. (2008a). *Test of language development-primary, Fourth edition*. Pro-Ed.

Hammill, D. D., & Newcomer, P. L. (2008b). *Test of language development-intermediate, Fourth edition*. Pro-Ed.

Hendricks, A. E., & Adlof, S. M. (2017). Language assessment with children who speak non-mainstream dialects: Examining the effects of scoring modifications in norm-referenced assessment. *Language, Speech, and Hearing Services in Schools*, *48*(3), 168-182.

Hirsch, E. D. (2003). Reading comprehension requires knowledge-of words and the world: Scientific insights into the fourth-grade slump and the nation's stagnant comprehension scores. *American Educator, 27(1)*, 10-22, 28-29, 48.

Iruka, I. U., Curenton, S. M., & Gardner, S. (2015). How changes in home and neighborhood environment factors are related to change in Black children's academic and social development from kindergarten to third grade. *The Journal of Negro Education*, *84*(3), 282-297.

Iruka, I. U., Curenton, S. M., Sims, J., Blitch, K. A., & Gardner, S. (2020). Factors associated with early school readiness profiles for Black girls. *Early Childhood Research Quarterly*, *51*, 215-228.

Kaufman, A. S., & Kaufman, N.L. (2004). *Kaufman brief intelligence test,* (2nd ed.) Pearson.

Kim, Y.-S. G. (2020). Hierarchical and dynamic relations of language and cognitive skills to reading comprehension: Testing the direct and indirect effects model of reading (DIER). *Journal of Educational Psychology, 112*(4), 667-684.

Koball, H., Moore, A, & Hernandez, J. (2021). *Basic facts about low-income children: Children*

under 9 years, 2019. National Center for Children in Poverty, Bank Street College of Education.

Ladson-Billings, G. (2009). *The dreamkeepers: Successful teachers of African American children*. John Wiley & Sons.

Language and Reading Research Consortium. (2015). The dimensionality of language ability in young children. *Child Development*, *86,* 1948–1965.

Language and Reading Research Consortium. (2017). Oral language and listening comprehension: Same or different constructs. *Journal of Speech, Language, and Hearing Research, 60*, 1273–1284.

Lindo, E. J. (2006). The African American presence in reading intervention experiments. *Remedial and Special Education*, *27*(3), 148–153.

Logan, J. A. R., Justice, L. M., Yumus, M., & Chaparro-Moreno, L. J. (2019). When children are not read to at home: The million word gap. *Journal of Developmental & Behavioral Pediatrics, 40 (5),* 383–386.

Mills, M. T. (2015). Narrative performance of gifted African American school-aged children from low-income backgrounds. *American Journal of Speech-Language Pathology*, *24*(1), 36-46.

Morgan, P. L., Farkas, G., Cook, M., Strassfeld, N. M., Hillemeier, M. M., Pun, W. H., & Schussler, D. L. (2017). Are Black children disproportionately overrepresented in special education? A best-evidence synthesis. *Exceptional Children, 83*, 181–198.

Murphy, K. A., LARRC, & Farquharson, K. (2016). Investigating profiles of lexical quality in preschool and their contribution to first grade reading. *Reading and Writing: An Interdisciplinary Journal, 29 (9)*, 1745–1770.

National Center for Education Statistics. (2019). *The nation's report card: Mathematics and reading assessments 2019.* Institute of Education Sciences, US Department of Education. https://nces.ed.gov/nationsreportcard/.

National Center for Education Statistics. (2022). *NAEP long-term trend trend assessment results: Reading and mathematics.* Institute of Education Sciences, US Department of Education. https://www.nationsreportcard.gov/highlights/ltt/2022/.

NICHD Early Child Care Research Network. (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology, 41*, 428–442.

Pearson, P. D., Palincsar, A. S., Biancarosa, G., & Berman, A. I. (2020). Reaping the rewards of the reading for understanding initiative. *National Academy of Education*.

Purpura, D. J. (2019). Language clearly matters; methods matter too. *Child Development, 90* (6), 1839–1846.

Rad, M. S., Martingano, A. J., Ginges, J. (2018). Toward a psychology of *Homo sapiens*: Making psychological science more representative of the human population. *Proceedings of the National Academy of Sciences, 115*, 11401–11405.

Reed, D. K., Sorrells, A. M., Cole, H. A., Takakawa, N. N. (2012). The ecological and population validity of reading interventions for adolescents: Can effectiveness be generalized? *Learning Disability Quarterly, 36(3*), 131-144.

Roberts, S. O., Bareket-Shavit, C., Dollins, F. A., Goldie, P. D., & Mortenson (2020). Racial inequity in psychological research: Trends of the past and recommendations for the future. *Perspectives on Psychological Science, 15* (6), 1–15.

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis) abilities:

Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research (97–110),* Guilford.

Semingson, P., & Kerns, W. (2021). Where is the evidence? Looking back to Jeanne Chall and enduring debates about the science of reading. *Reading Research Quarterly*, *56*, S157-S169.

Seymour, H.N., Roeper, T., & de Villiers, J. (2003). *Diagnostic Evaluation of Language Variation* (*DELV) Screening Test*. The Psychological Corporation.

Shi, L., & Stevens, G. D. (2021). *Vulnerable populations in the United States* (3rd Ed.). Wiley.

Skiba R. J., Artiles, A. J., Kozleski, E. B., Losen, D. J., Harry, E. (2016). Risks and consequences of oversimplifying educational inequities: A response to Morgan et al. (2015). *Educational Researcher, 45(3)*, 221–225.

Smedley, A., & Smedley, B. D. (2005). Race as biology is fiction, racism as a social problem is real: Anthropological and historical perspectives on the social construction of race. *American Psychologist, 60*(1), 16–26.

Snow, C. E., Barnes, W. S., Chandler, J., Goodman, I. F., & Hemphill, L. (2013). *Unfulfilled expectations*. Harvard University Press.

Snow, C.E., Burns, M.S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. National Academy Press.

Stockard, J. (2010). Promoting reading achievement and countering the "fourth-grade slump": The impact of direct instruction on reading achievement in fifth grade. *Journal of Education for Students Placed at Risk*, *15*(3) 218–240.

Terry N. P., Connor C. M., Petscher Y., & Conlin C. R. (2012). Dialect variation and reading: Is change in nonmainstream American English use related to reading achievement in first

and second grades? *Journal of Speech, Language, and Hearing Research*, 55, 55–69.

Terry, N. P., Gatlin, B., & Johnson, L. (2018). Same or different: How bilingual readers can help us understand bidialectal readers. *Topics in Language Disorders*, *38(1),* 50–65.

Terry, N. P., Gerido, L. H., Norris, C. U., Johnson, L. R., & Little, C. (2022). Building a framework to understand and address vulnerability to reading difficulties among children in schools in the United States. *New Directions for Child and Adolescent Development*, 2022(183–184), 9–26.

US Census Bureau. (2015). *2014 American Community Survey*. https://www.census.gov/programs-surveys/acs/data/data-tables.html.

Wagner, R., Torgeson, J., Rashotte, C., & Pearson, N. A. (2013). *Comprehensive test of phonological processing* (2nd *ed.).* Pro-Ed.

Washington, J. A., Branum-Martin, L., Lee-James, R., & Sun, C. (2019). Reading and language performance of low-income, African American boys in grades 1–5. *Reading & Writing Quarterly*, *35*(1), 42–64.

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. *Journal of Speech, Language, and Hearing Research, 49(2),* 232–247.

Wiederholt, J., & Bryant, B. (2012). *Gray oral reading tests* (5th Ed.). Pro-Ed.

Woodcock , R., Mather, N., & Schrank, F. A. (2004). *Woodcock diagnostic reading battery* (3rd ed.). Riverside.

Table 1.

Demographic Characteristics of the Participants by each Grade

Grade (%)	*n*	Female	Age in Years		KBIT		DVAR	
			M	*SD*	*M*	*SD*	*M*	*SD*
1	228	52.4	6.8	0.4	95.9	13.8	64.0	29.7
2	189	56.9	7.9	0.5	97.2	14.8	54.7	29.3
3	193	51.8	8.8	0.5	99.6	13.6	49.0	24.9
4	147	49.3	9.9	0.5	97.3	12.3	49.4	27.7

Note. KBIT = the Kaufman Brief Intelligence Test, DAVR = dialect variation score computed from Part I of the Diagnostic Evaluation of Language Variation Screening Test.

Table 2

Mean Performance by Grade Level and Results of Independent t-tests comparing Grade 1 and Grade 4 for Each Measure

Measure		Grade 1	Grade 2	Grade 3	Grade 4	*t*-value (*df*)	Cohen's *d*
WJ3 Letter-Word ID	*M* (*SD*)	109.7 (14.0)	103.3(13.9)	100.5 (11.5)	97.9 (12.6)	8.23 (369)	13.47
	Range	67-143	60-145	59-131	60-138		
	% -1SD	6%	10%	8%	11%		
WJ3 Word Attack	*M* (*SD*)	111.3 (14.6)	103.0 (11.8)	100.7 (10.5)	99.1 (11.7)	8.81 (352.7)*	13.56
	Range	64-139	69-148	76-136	56-127		
	% -1SD	7%	4%	7%	11%		
WJ3 Reading Fluency	*M* (*SD*)	108.6 (11.5)	101.4 (11.9)	97.2 (9.7)	94.9 (11.4)	10.23 (290)	11.46
	Range	74-153	67-148	69-122	66-132		
	% -1SD	1%	11%	12%	21%		
WJ3 Passage Comprehension	*M* (*SD*)	102.2 (13.9)	98.1 (11.0)	93.8 (7.4)	89.7 (8.5)	10.73 (367.0)*	12.09
	Range	60-134	70-132	67-122	57-115		
	% -1SD	11%	12%	12%	26%		
GORT Oral Reading Index	*M* (*SD*)	88.6 (13.1)	90.2 (11.9)	87.1(10.4)	82.9 (10.9)	4.54 (340.3)*	12.25
	Range	57-126	54-123	57-115	52-113		
	% -1SD	35%	30%	44%	57%		
CTOPP Blending Words	*M* (*SD*)	7.8 (2.6)	6.9 (2.3)	6.1 (1.9)	6.2 (2.1)	6.27 (367)	2.4
	Range	1-17	1-17	1-13	1-13		
	% -1SD	44%	68%	80%	77%		

Measure		Grade 1	Grade 2	Grade 3	Grade 4	*t*-value (*df*)	Cohen's *d*
CTOPP Elision	*M* (*SD*)	8.6 (2.5)	8.1 (2.4)	7.7 (2.5)	7.2 (2.8)	4.94 (277.3)*	2.59
	Range	1-16	3-16	2-15	1-14		
	% -1SD	27%	50%	63%	59%		
TOLD Morphology	*M* (*SD*)	8.2 (3.0)	8.4 (2.9)	7.6 (2.5)	7.1 (2.5)	3.86 (347.3)*	2.78
	Range	2-16	2-16	3-16	3-16		
	% -1SD	43%	41%	59%	62%		
TOLD Picture Vocabulary	*M* (*SD*)	9.9 (2.8)	9.5 (2.6)	8.1 (2.6)	7.4 (2.4)	8.87 (369)	2.62
	Range	1 – 16	4-17	1-16	2-16		
	% -1SD	20%	26%	40%	57%		
TOLD Syntax	*M (SD*)	9.7 (2.4)	9.0 (3.0)	7.5 (2.8)	6.9 (2.8)	9.88 (275.4)*	2.53
	Range	3-15	5-15	2-16	2-14		
	% -1SD	18%	33%	60%	62%		

Note. M = test mean. *SD* = Standard deviations. % -1SD = the percentage of students performing at or below one standard deviation from the test mean. Range = lowest to highest standard score in the sample. All t-tests were significant at the p-value less than .001. Type I error adjustment with Bonferroni correction, p-value less than .005 was considered as significant. * t-values with "equal variance not assumed".

Table 3

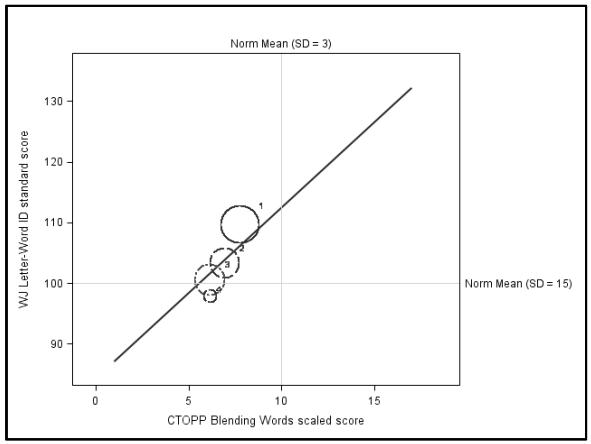
Partial Correlations between Performance on Reading and Oral Language Measures (controlling for age).

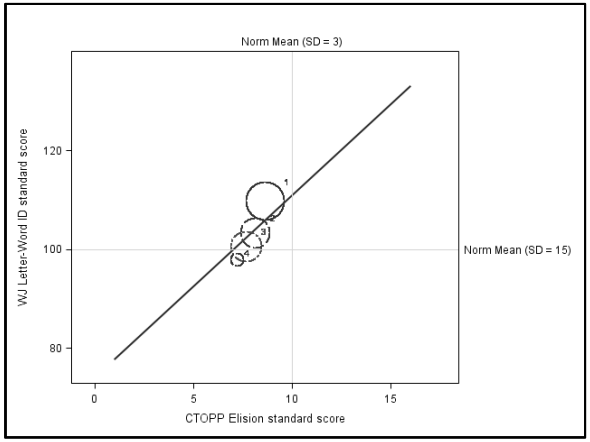
Measure	1	2	3	4	5	6	7	8	9	10
1. WJ3 Letter-Word ID	—									
2. WJ3 Word Attack	.87	—								
3. WJ3 Reading Fluency	.80	.74	—							
4. WJ3 Passage Comprehension	.85	.78	.76	—						
5. GORT Oral Reading Index	.79	.69	.79	.77	—					
6. CTOPP Blending Words	.48	.50	.41	.47	.40	—				
7. CTOPP Elision	.68	.67	.57	.63	.60	.46	—			
8. TOLD Morphology	.44	.42	.40	.47	.49	.40	.46	—		
9. TOLD Picture Vocabulary	.49	.44	.48	.55	.51	.41	.43	.58	—	
10. TOLD Syntax	.54	.52	.53	.58	.50	.42	.51	.47	.53	—

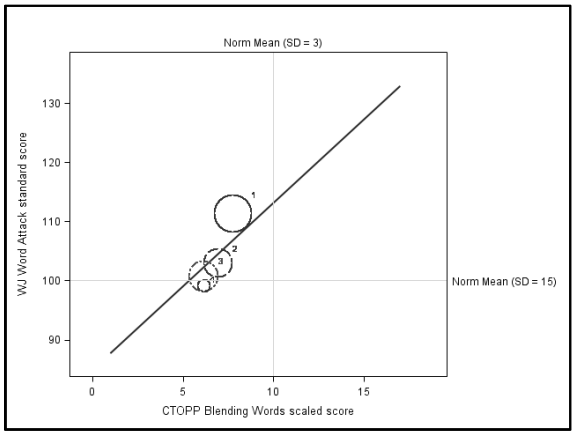
Note: all correlations are significant at $p < .001$.

Figure 1

Visual representations of the relations between WJ-Letter Word Identification (top row) and Word Attack (bottom row) scores and CTOPP-Blending Words and Elision Scores across Grades 1 and 4.







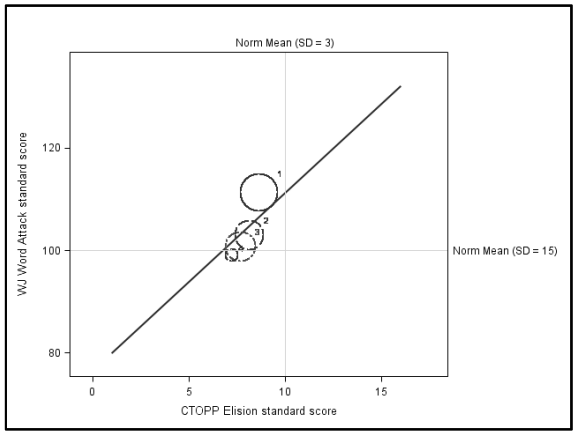


Figure 2

Visual representations of the relations between WJ-Reading Fluency scores and CTOPP-Blending Words and Elision Scores across Grades 1 and 4.

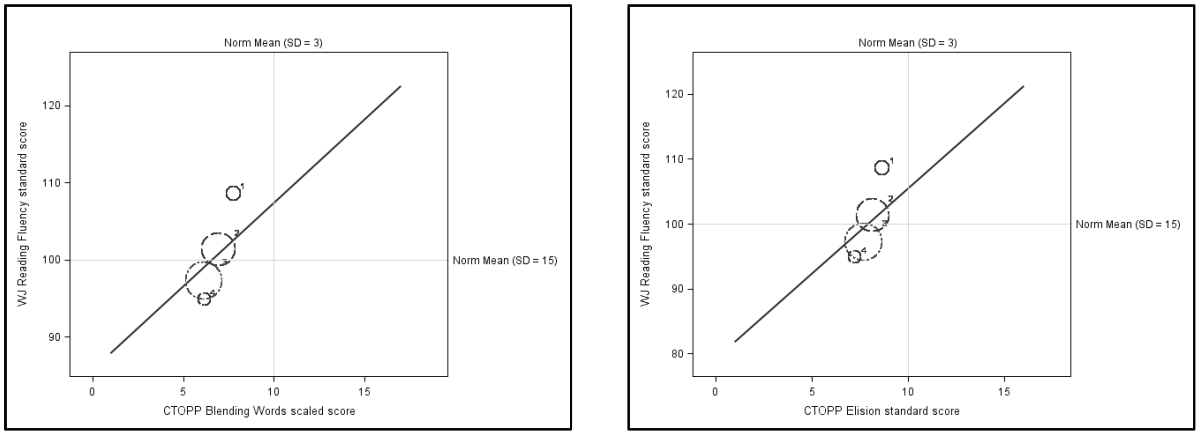


Figure 3

Visual representations of the relations between WJ-Passage Comprehension scores (top row) and GORT Sum of Scaled Scores (bottom row) and TOLD scores across Grades 1 and 4.