Prekindergarten and Kindergarten Predictors of Reading Comprehension in Monolingual English and Spanish-English Bilingual 6th Graders

Shelley Gray[a*], Robert Davies[b], Marilyn Thompson[a], Maria Adelaida Restrepo[d], Mindy Bridges[c], Savannah Romeo[a], Jinxiang Hu[c], and Kate Cain[b]

Paper accepted for publication Journal of Speech, Language, and Hearing Research on 1/04/2026.

[a*]Shelley Gray, College of Health Solutions, Arizona State University, US, Shelley.Gray@asu.edu

[b]Robert Davies, Department of Psychology, Lancaster University, UK, r.davies1@lancaster.ac.uk

[a]Marilyn S. Thompson, The Sanford School, Arizona State University, US, M.Thompson@asu.edu (now at Oregon State University Marilyn.Thompson@oregonstate.edu)

[d]Maria Adelaida Restrepo, College of Health Solutions, Arizona State University, US. Now at Department of Communication Sciences and Disorders, University of South Florida, US, marestrepo@usf.edu

[c]Mindy Bridges, University of Kansas Medical Center, US, mbridges2@kumc.edu

[a]Savannah Romeo, College of Health Solutions, Arizona State University, US, savannah.romeo@asu.edu

[c]Jinxiang Hu, University of Kansas Medical Center, US, jhu2@kumc.edu

[b]Kate Cain, Department of Psychology, Lancaster University, UK, k.cain@lancaster.ac.uk

Keywords:

reading comprehension; children; predictors

Preregistration

We completed an initial pre-registration of the analysis of Grade 6 reading comprehension in monolingual English and Spanish-English bilingual children (referred to as Dual Language Learners in that document). Prior to data analyses we revised our pre-registration plan. The decision to use item-level Bayesian mixed-effects models was specified in the revised preregistration prior to analysis. Links to preregistration documents include: https://osf.io/4sfnu, https://osf.io/whacq, and https://osf.io/v9ntx and the addendum https://osf.io/c2pwz. The revised pre-registration plan that we adhered to in these analyses is here: https://osf.io/dt5gu.

Conflict of Interest

The authors declare there are no real or potential conflicts of interest that could be seen as having an influence on this research.

Abstract

Purpose: Poor reading comprehension performance by children in the US is a continuing concern. Early identification and intervention can reduce the number of children with significant reading problems. We documented prekindergarten and kindergarten predictors of Grade 6 reading comprehension for English monolingual and Spanish-English bilingual groups and identified measures available to educators to flag children at risk for future reading comprehension problems.

Method: In Grade 6, children in the monolingual (N = 88) and bilingual groups (N = 95) completed a reading comprehension measure. These children were in a longitudinal study with previously completed code-related, vocabulary, grammar, listening comprehension, higher-level language, working memory, and nonverbal IQ measures in prekindergarten and kindergarten.

Results: Data were analyzed separately by language group. We fit a series of Bayesian mixed-effects models incorporating prekindergarten and kindergarten measures as predictors. For the monolingual group, the most promising prekindergarten predictor was letter identification, and the most promising kindergarten predictors were vocabulary, grammar/morphology, and listening comprehension. For the bilingual group, the most promising prekindergarten predictors were English vocabulary, listening comprehension in Spanish, and memory updating in Spanish, and the most promising kindergarten predictor was English vocabulary.

Conclusion: We suggest measures that could be administered in prekindergarten and kindergarten to flag students who may be at risk for future reading comprehension problems. We review other steps that schools and families may take to prevent reading comprehension problems.

Introduction

Given the relatively unchanged National Assessment of Educational Progress (NAEP) reading scores over the past two decades, it is fair to ask whether anything can be done to improve reading comprehension in the U.S., especially for bilingual children who, as a group, score significantly lower than their monolingual English peers (e.g., Friesen et al., 2022; Proctor et al., 2012; Silverman et al., 2015). Prevention is one potential solution (McCardle et al., 2001; Pikulski, 1994; Snow et al., 1998), with convincing evidence that early identification and intervention can reduce the number of children with significant reading problems (Linan-Thompson et al., 2006; McMaster et al., 2005). Because Head Start offers prekindergarten (PK) screening and 73% of public schools offer kindergarten (K) screening (Shields et al., 2016), accurate early identification is a promising approach for identifying children who need tailored instruction early in schooling; however, the best combination of predictors for later reading comprehension has not yet been identified for either monolingual or bilingual children.

Research shows that reading comprehension is a complex skill set involving multiple constructs (e.g., LARRC & Logan, 2017; Oakhill & Cain, 2012; Ortiz et al., 2012). To identify early indicators of later reading comprehension, it is necessary to assess a range of theoretically motivated constructs in early childhood. For three decades the Simple View of Reading (Gough & Tunmer, 1986) has provided a valuable, testable, framework for understanding reading comprehension that has been supported in English (e.g., Hoover & Gough, 1990; Kershaw & Schatschneider, 2012; LARRC, 2015c; Oakhill et al., 2003; Tilstra et al., 2009) and many other languages including Spanish (e.g., Mancilla-Martinez & Lesaux, 2010) and Greek (Protopapas et al., 2012). In general, word reading and linguistic comprehension, the two central components of the framework, have explained from 40% to 80% of the variance in reading comprehension in

English-speaking children (Bell & Wheldall, 2022; Joshi et al., 2015; Nation, 2019), with the majority of these studies being conducted in grades 1-3.

Only a small number of longitudinal studies have examined predictors of reading comprehension that extend beyond the code-related and oral language skills emphasized in the Simple View of Reading (Gough & Tunmer, 1986) by incorporating additional cognitive factors such as working memory and nonverbal IQ. In these studies, “code-related skills” included foundational processes that support word-reading development, including phonological awareness, letter identification, decoding, and orthographic knowledge. To provide the most appropriate points of comparison for our work, we reviewed studies in which children were assessed beginning in K or first grade, and in which those early skills were used to predict reading comprehension outcomes measured from Grade 2 through high school. Apart from LARRC and Chiu (2018) there were no longitudinal studies that assessed children in preschool to predict later reading comprehension. We included K and grade 1 as predictor grades when children are first learning to read to support early screening for language and literacy disorders.

Two studies included poor readers (Fuchs et al., 2012; Peng et al., 2019). Each study measured code-related skills and oral language plus cognitive abilities such as non-verbal reasoning, verbal memory, working memory, rapid naming, or IQ in English. Across all of these studies code-related skills, oral language, and cognitive measures were significant predictors of subsequent reading comprehension from one to nine years later. These results suggest that to best predict reading comprehension, cognitive as well as code-related and oral language measures should be assessed in early grades.

Across reviewed studies participants were from the US, UK, Norway, Finland, and Canada meaning that they spoke multiple languages and represented a wide variety of school systems.

We know that bilingual and monolingual children differ in language experience, vocabulary, and developmental timing; however, prior work suggests that many core processes underlying reading comprehension operate similarly within each group. For example, studies of Spanish–English bilingual learners have shown that vocabulary, syntax, and word-reading skills predict English reading comprehension, and that Spanish language skills can support English comprehension through cross-linguistic transfer (Farnia & Geva, 2013; Proctor et al., 2006; Mancilla-Martinez & Lesaux, 2010). However, it is also well established that bilingual and monolingual children do not share identical underlying measurement structures for oral language or reading comprehension. Studies conducted with the same or overlapping longitudinal samples used in the present study have demonstrated that oral language skills, including vocabulary, grammar, and listening comprehension, are not measurement-invariant across Spanish–English bilingual and English monolingual children (LARRC, Yeomans-Maldonado, Bengochea, & Mesa, 2018). This lack of invariance makes direct statistical comparisons across the two language groups inappropriate. Accordingly, our aim is not to compare predictors between bilingual and monolingual children, but rather to identify, within each group separately, the early language, code-related, and cognitive skills most strongly associated with Grade 6 reading comprehension. This approach permits careful interpretation of group-specific developmental patterns while acknowledging that similar processes may underlie reading development even when constructs do not behave identically across populations.

In these longitudinal studies, significant code-related predictors of later reading comprehension included letter knowledge (Manu et al., 2021 in Finnish monolinguals; Peng et al., 2019 in English monolinguals; Psyridou et al., 2021 in Finnish monolinguals), phonological awareness (Farnia & Geva, 2013 in English bilinguals; Psyridou et al., 2021 in Finnish monolinguals), word reading (Farnia & Geva, 2013 in English bilinguals; Fuchs et al., 2012 in English monolinguals; Lervag et al., 2019 in Norwegian monolinguals; Peng et al., 2019 in English monolinguals; Psyridou et al., 2021 in Finnish monolinguals; Sparks et al., 2014 in English monolinguals), and reading fluency (Fuchs et al., 2012 in English monolinguals).

Oral language skills were also significant predictors of later reading comprehension using an oral language factor that included vocabulary, grammar, and discourse (LARRC & Chiu, 2018) as well as separate indicators of grammar (Lervag et al., 2018), syntax (Farnia & Geva, 2013), vocabulary (Farnia & Geva, 2013; Lervag et al., 2019; Manu et al., 2021; Peng et al., 2019; Psyridou et al., 2021; Sparks et al., 2014), reading inferencing (Lervag et al., 2019; Oakhill & Cain, 2012), listening comprehension (Fuchs et al., 2012; Lervag et al., 2019; Manu et al., 2021), comprehension monitoring (Oakhill & Cain, 2012), and passage comprehension (Sparks et al., 2014). Of these, vocabulary was a significant predictor in 6 of 7 studies with reading comprehension measured as early as second grade and as late as high school.

The longitudinal studies assessed cognitive development using a wide range of measures. Significant predictors of later reading comprehension included nonverbal reasoning (Fuchs et al., 2012; Peng et al, 2019), verbal memory (Lervag et al., 2018), rapid naming (Farnia & Geva, 2013; Fuchs et al., 2012; Manu et al., 2021; Psyridou et al., 2021), working memory (Farnia & Geva, 2013) and IQ (Sparks et al., 2014).

Together, this literature suggests that while many early skills relate to later reading comprehension, fewer studies have examined how predictors operate across adjacent early grades within linguistically distinct groups. The present study addresses this gap by examining PK and K predictors within monolingual and bilingual groups separately.

Assessments in the present longitudinal study included multiple code-related, oral language, and cognitive measures for two groups of children: those who entered PK speaking English and those who entered PK speaking primarily Spanish. As described below, those who entered PK speaking primarily Spanish became bilingual as they were educated in English only schools beginning in K.

Although the present study was not designed to evaluate diagnostic classification or screening accuracy, identifying early measures associated with later reading comprehension is important because it informs hypotheses about which constructs warrant attention in early educational contexts. In seeking to make accurate and generalizable predictions of later reading comprehension, our analyses targeted the accuracy of response to comprehension test items at Grade 6 so that prediction models could take into account both participant- and item-level sources of variance.

Our research questions were:

1. Which language, cognitive and code-related skills in PK or K best predict reading comprehension in Grade 6 for the monolingual and bilingual groups?
2. Of these predictors, which PK and K measures could flag children who may benefit from more intense instruction to promote future reading comprehension?

Method

This study was approved by the Arizona State University Internal Review Board under study numbers 1010005588 and 00006696.

Participants

Participants were initially enrolled in a 5-year longitudinal study conducted by LARRC. The children who began PK speaking English (monolingual group; $N = 397$) were from Arizona, Kansas, Missouri, Nebraska, and Ohio. The children who began PK speaking Spanish (bilingual group; $N = 280$) were from Arizona. In Arizona by state law, instruction must be provided in English.

Participants were recruited through partnerships with multiple school districts and individual schools selected based on student population size, demographic diversity, and willingness to participate in a longitudinal study. Within participating schools, all PK teachers were invited to participate, and children were recruited through those classrooms. Families of all children in participating classrooms received study information and consent materials. Children were not randomly sampled; instead, eligibility was determined through parent and teacher screening to confirm that participants met inclusionary criteria and were able to complete the study assessments.

Because children entered the study in PK and then matriculated into many different kindergartens, the number of schools represented increased. Additional school changes occurred as families relocated during grades K–5, resulting in a large number of schools represented in Grade 6. In PK monolingual participants were enrolled in 25 different schools, in K 58 different schools, and in Grade 6 in 57 different schools. In PK bilingual participants were enrolled in 33 different schools, in K in 41 different schools, and in Grade 6 in 54 different schools. Nine or fewer participants were enrolled in the same school at any grade level for either group.

The monolingual English-speaking sample was drawn from multiple states, increasing geographic and educational variability and strengthening the generalizability of findings within that group. In contrast, the Spanish–English bilingual sample was recruited from a single state, which limited geographic variability but resulted in a more linguistically cohesive group of children who primarily spoke Mexican Spanish. This sampling structure aligns with our analytic approach of examining predictors within each language group rather than making direct comparisons across groups.

The purpose of both the LARRC and MBRC studies was to identify and model oral language, cognitive, environmental, and code-related skills important for reading comprehension. In the original LARRC study, data were collected annually from PK through Grade 3. Subsequently, children were reconsented and enrolled in a follow-up longitudinal study conducted by the Monolingual and Bilingual Reading Comprehension (MBRC) group, which collected additional data in Grade 6. Because not all original participants were retained through Grade 6, we conducted attrition analyses to determine whether children who remained in the study differed systematically from those who were lost to follow-up. Among monolingual participants, attrition was not associated with gender, ethnicity, race, income, or eligibility for free or reduced-price lunch, although attrition varied by site, with participants from Nebraska more likely to be lost to follow-up than those from Arizona. Among bilingual participants, no significant demographic differences were observed between retained and attrited participants. Accordingly, the analytic sample used in the present study reflects children with Grade 6 outcome data, and results should be interpreted as characterizing associations within the retained sample.

We also assessed whether the Ohio sample, which was lost during the transition from LARRC to MBRC, differed from other sites at PK. Results indicated that Ohio participants were more likely to be from low- and middle-income households and to qualify for free/reduced lunch, although no differences were found in other demographics.

Previous published studies involving the same LARRC and MBRC samples include: Currie et al., 2021; LARRC, 2015a; 2015b; 2015c; 2017a; 2017b; 2017c; Murphy et al., 2016; LARRC & McIlraith, 2018; LARRC et al., 2018; LARRC & Muijselaar, 2018; LARRC & Chiu, 2018; LARRC et al., 2018a; LARRC et al. 2018b; LARRC et al., 2019a; LARRC et al., 2019b; LARRC et al. 2021). These papers report analyses from grades PK through Grade 3 but have no overlap in research questions with the current study.

Inclusionary criteria for participants in their PK year included: (a) child spoke English as their native language for the monolingual group or Spanish as their native language for the bilingual group based on parent report (some researchers refer to these children as dual-language learners); (b) child had no significant speech, language, cognitive, sensory, or motor disabilities that would preclude participation in assessments according to parent and teacher reports; (c) child was attending PK, and (d) child was eligible to enter K the following year. Language dominance was not used to form subgroups within the bilingual sample. Instead, variability in Spanish and English skills was treated as continuous and reflected directly in the predictor measures used in the analyses.

For this study we used longitudinal data from PK, K, and Grade 6. Demographic information about the Grade 6 samples is presented in Table 1 for the monolingual group and Table 2 for the bilingual group. We intentionally separated the groups descriptively and in all analyses (LARRC, 2015a).

Procedures

Children were assessed in the spring of each school year in a quiet room at their school or at a community center or library. Data collection followed the same procedure each year. Trained research assistants (RAs) administered the measures 1:1. For the bilingual group, Spanish-English bilingual RAs provided the assessment in the target language with Spanish measures administered on a different day than English measures. Children completed about six hours of assessments divided into multiple sessions of 1 – 2 hours each. The order of test/task administration was counterbalanced across participants, although the measures were arranged in blocks based on the language of the day, the need for recording, and the age of the child. In PK and K children received $40 in literacy materials and parents received a $15 gift card as an incentive. In Grade 6 children received a $125 incentive.

Measures

Our reading comprehension outcome measure in Grade 6 was the Woodcock Reading Mastery Test - Passage Comprehension subtest (Woodcock, 1998), administered in English for both groups. This subtest required students to read a sentence or short passage and identify a missing word. To answer correctly, students must use a variety of comprehension and vocabulary skills to understand the passage.

The LARRC and MBRC studies included multiple measures of key constructs associated with reading comprehension. A subset of these measures (Table 3 for the monolingual group; Table 4 for the bilingual group) was selected based on their availability at both PK and kindergarten. This approach ensured consistency in construct representation across early grades and allowed us to examine whether the same skills, measured with comparable instruments, were associated with later reading comprehension across developmental stages. It also aligns with our goal of identifying screening measures that could be used flexibly in either PK or kindergarten settings.

We used measures administered in English for the monolingual group and in Spanish for the bilingual group. In PK the children came to school speaking primarily Spanish, thus it would not be appropriate to assess their skills in English. The exceptions were that we administered the English language Peabody Picture Vocabulary Test-4 (PPVT-4; Dunn & Dunn, 2007) and the English language Woodcock Reading Mastery Test (WRMT) Letter Identification (LI) task to children in the bilingual group because they were receiving school instruction in English. More details on measure selection are provided in the Analytic Approach section. Reliability calculations for measures administered to the monolingual group are reported in Supplementary Table 1 and to the bilingual group in Supplementary Table 2. Monolingual alphas in PK ranged from .71-.96 and in kindergarten from .57-.96. Bilingual group alphas ranged from .60-.97 in PK and .64-.96 in kindergarten.

Code-Related Skills in PK and K

In PK and K, the Woodcock Reading Mastery Test - *Letter ID* subtest (Woodcock, 1998) assessed the child's ability to identify alphabet letters presented in isolation. For the present analyses, we included scores from the English version at PK and K for both groups.

Oral Language

As shown in Tables 3 and 4, we assessed four oral language constructs: vocabulary, grammar, listening comprehension, and higher-level language. Although the CELF-4 and CELF-4 Spanish are designed for children ages 5-21, we administered them to four-year-olds in this study but used only raw scores for analyses.

We assessed vocabulary breadth in the monolingual English group using the Peabody Picture Vocabulary Test – 4th Edition (PPVT-4; Dunn & Dunn, 2007) and vocabulary depth using the Clinical Evaluation of Language Fundamentals – 4th Edition Word Classes subtest (CELF-4; Semel, Wiig, & Secord, 2003). Both are receptive measures. The bilingual group completed the Clinical Evaluation of Language Fundamentals – 4th Edition Spanish Edition Word Classes subtest (Semel, Wiig, & Secord, 2003) in Spanish in PK and K and the PPVT-4 in English in K.

For grammar we administered the CELF-4 Word Structure subtest to the monolingual group and the CELF-4 Spanish Word Structure subtest to the bilingual group. Both assess syntax and morphology in their respective languages.

In PK and K, we assessed listening comprehension using modifications of the CELF-4 Understanding Spoken Paragraphs subtest for the monolingual group and the CELF-4 Spanish Understanding Spoken Paragraphs for the bilingual group. The bilingual group also completed the English CELF-4 Understanding Spoken paragraphs in K. Because the CELF-4 and CELF-4 Spanish do not have passages for PK, we created new paragraphs and questions using lexile and readability measures appropriate for PK. In K the existing paragraphs were divided so that only 1 trial paragraph and two additional paragraphs were administered.

To assess higher-level language, we administered the Inference-Making Stories Task, an experimenter-made measure based on the work of Cain and Oakhill (1999), Oakhill and Cain, (2012), and Currie and Cain (2015). Stories were presented in English to the monolingual group and Spanish to the bilingual group in both PK and K. Children listened to an audio recording of two short narratives and answered eight open-ended questions after each passage. In K one of the stories was new and one story had been administered previously in PK. For each passage, four questions assessed the ability to make local coherence inferences and four assessed the ability to make global coherence inferences.

We also administered the Picture Arrangement Task (PAT), adapted from the Picture Arrangement Test of the Wechsler Intelligence Scale for Children-3 (WISC-III UK edition; Wechsler, 1992) that assessed children's ability to sequence a series of picture cards into a coherent story. The total number of correct stories was tallied for the raw score, which was used in analyses (possible range: 0–12). Internal consistency for our sample was $\alpha = .85$.

Working Memory

Working memory was assessed using the Auditory Working Memory subtest of the Woodcock-Johnson Test of Cognitive Ability III normative update (WJ-III, Woodcock, 1998) for the monolingual group and the Auditory Working Memory subtest of the Bateria III Woodcock-Muñoz (Muñoz-Sandoval et al., 2005) for the bilingual group. This assessment tests children's ability to listen to a list of words and numbers then to name the words in order followed by the numbers in order. The Spanish version does not include monosyllabic words.

We also administered an experimenter-made working memory updating measure based on the work of Belacchi et al. (2010) that evaluated the child's ability to modify the contents of working memory using comparison of objects. The child saw pictures of an increasing number of objects and was asked a question like, "Tell me the names of the one/two/three/four/five smallest things." Instructions and prompts were in Spanish for the bilingual group.

Nonverbal IQ

To assess nonverbal IQ, we administered the Kaufman Brief Intelligence Test – 2nd Edition Matrices subtest (Kaufman & Kaufman, 1997) in PK, but not in K. The Matrices subtest assesses children's ability to solve new problems, perceive relationships and complete visual analogies without testing vocabulary or language skills. Spanish prompts were provided for the bilingual sample. This is in line with the test manual instructions, which state, “When assessing individuals whose comprehension of spoken English is limited, such as examinees who speak little English or those with hearing impairments, the examiner may present directions for Matrices in any foreign language or alternative communication method (such as signing) shared by the examinee and the examiner. For the convenience of Spanish-speaking examiners, each Matrices easel page presents administration instructions in both English and Spanish.”

Results

Analysis Method

We analyzed Grade 6 reading comprehension using item-level response accuracy on the WRMT Passage Comprehension subtest. Building on prior work using item-level approaches to reading comprehension (e.g., Kulesz et al., 2016; Miller et al., 2014; Spencer et al., 2019), we used Bayesian generalized linear mixed-effects models (GLMMs) to estimate how prekindergarten and kindergarten language, cognitive, and code-related measures predict the log-odds of a correct response. In these models, the fixed effects represent the effects of early predictors on comprehension accuracy, while the random effects account for variability associated with differences between children and differences between items. The fixed effects in these models represent the effects of early language, cognitive, and code-related predictors on comprehension accuracy, while the random effects account for variability associated with differences between children and differences between items. Modeling at the item level allows simultaneous estimation of between-child differences in overall comprehension ability and between-item differences in difficulty, preserving variability that would be obscured by summed scores and aligning with prior longitudinal prediction work (e.g., Kulesz et al., 2016; Miller et al., 2014; Spencer et al., 2019).

We assumed a Bernoulli likelihood function for item-level response accuracy and specified weakly informative prior distributions for population- and group-level parameters to stabilize estimation and regularize extreme estimates while allowing the data to primarily determine parameter values. These priors reflect the assumption that fixed effects are likely to fall within a moderate range on the log-odds scale and that random effects variances are positive and typically small (McElreath, 2020). For the primary models, we specified Normal (mean = 0, sd = 1) priors for fixed effects, Half-Student's t (location = 0, scale = 10, shape = 3) priors for random effects variances, and LKJ (shape = 2) priors for random effects covariance structures. The models included child-level and item-level random intercepts, along with random slopes where appropriate, to reflect the hierarchical structure of the data and to allow the effects of early predictors to vary across children and across items. This specification reflects the expectation that both overall comprehension ability and the influence of early skills may differ across items of varying difficulty. Residual item-level variance remained substantively non-zero after accounting for child-level differences, indicating that responses varied across items beyond what would be expected if Rasch assumptions fully held in this sample. Accordingly, modeling item-level accuracy with item random effects provides a more appropriate framework than total-score approaches for estimating associations between early predictors and later reading comprehension.

To address our research questions, we fitted a series of models. To evaluate which language, cognitive, or code-related skills influenced reading comprehension in Grade 6 we fitted: (1) models including (as predictors or fixed effects) PK measurements of grammar, vocabulary, higher-level language, listening comprehension, nonverbal IQ, and working memory; and (2) models including K measurements of the same constructs. PK and K predictors were modeled separately to (a) avoid collinearity introduced by including multiple timepoints in a single model and (b) evaluate the unique predictive contribution of skills assessed at each developmental stage. Analyses were conducted separately within each language group given prior evidence of measurement non-invariance across monolingual and bilingual samples (LARRC, Yeomans-Maldonado, Bengochea, & Mesa, 2018).

Data Analysis Results

The initial LARRC dataset included measures administered in PK (N = 397 monolingual, 280 bilingual) and K (N = 373 monolingual, 263 bilingual). Analyses were conducted using participants with Grade 6 WRMT Passage Comprehension outcome data (N = 88 monolingual, 95 bilinguals; see Participants for attrition details). Results are presented separately for the monolingual and bilingual groups, with PK and K predictor models reported within each group.

Using Bayesian mixed-effects models allowed us to include all available item-level responses without listwise deletion; however, this analytic approach was selected primarily to appropriately represent child- and item-level sources of variance rather than to address missing data. The presence of missing item-level data did not drive the choice of analytic approach; rather, item-level modeling was selected to appropriately represent child- and item-level sources of variance and to address limitations of total-score approaches. Given some attrition, results should be interpreted as estimating associations within the retained analytic samples, rather than as definitive population-level effects.

Model Estimates for the Monolingual Group

Table 5 presents estimates from the models used to analyze monolingual group responses, including (1) PK predictor models and (2) K predictor models. These analyses address Research Question 1 by identifying which PK and K predictors are most strongly associated with later reading comprehension within the monolingual group. For each model we report coefficient estimates on the log-odds scale, along with standard errors and 95% credible interval bounds. Posterior probabilities are interpreted descriptively, with values $\geq .95$ indicating strong evidence for an effect and values between .66 and .95 indicating moderate evidence. Posterior distributions for each predictor are also displayed in Figures 2–3.

For the monolingual group, PK letter identification scores were associated with a higher probability of correct responses in Grade 6 reading comprehension. The posterior distribution indicates strong evidence for a positive effect ($\beta = .39$), with approximately 95% of the distribution lying above zero. The credible interval suggests that the effect of PK code-related skills ranges from small ($\beta \approx .03$) to relatively large ($\beta \approx .77$), while the probability of a negative effect is very small (< 2.5%). There was also moderate evidence that higher PK scores on auditory memory, nonverbal intelligence (KBIT-2 Matrices), and grammatical knowledge (CELF-4 Word Structure) were associated with a higher probability of correct responses.

For the kindergarten models, higher vocabulary (PPVT-4) and grammatical knowledge (CELF-4 Word Structure) scores were associated with a higher probability of correct responses in Grade 6 reading comprehension. The posterior distributions provide strong evidence for positive effects of these predictors. There was also moderate evidence that listening comprehension (CELF-4 Understanding Spoken Paragraphs) was positively associated with response accuracy.

Figures 2 and 3 display the posterior probability distributions for the PK and K predictors. Density plots illustrate the distribution of estimated effects for each predictor, while the dot-and-whisker plots summarize these distributions using posterior medians and credible intervals.

Model Estimates for the Bilingual Group

We present estimates for each model used to analyze bilingual group responses in Table 6 including: (1) PK predictor model estimates and (2) K predictor model estimates. We also present these estimates as posterior probability distributions (Figures 5-7). These analyses address Research Question 1 within the bilingual group by identifying early predictors probabilistically associated with later English reading comprehension.

Together, the results summarized across PK and K models for each group address Research Question 2 by identifying early measures that show the strongest probabilistic associations with later reading comprehension, while explicitly acknowledging uncertainty and sample-specific limitations.

For the bilingual group, higher PK scores on memory updating in Spanish, vocabulary in English (PPVT-4), and listening comprehension in Spanish (CELF-4 Understanding Spoken Paragraphs) were associated with a higher probability of correct responses in Grade 6 reading comprehension. The data suggest that higher PK scores on the Spanish Inference Making Stories test of higher-level language are also associated with response accuracy, although the posterior distributions are compatible with small, null, or negative as well as positive effects.

In comparison, only higher kindergarten vocabulary scores (PPVT-4) showed strong evidence (posterior probability ≥ 0.95) of predicting a higher probability of correct responses in Grade 6 reading comprehension. There was also moderate posterior probability (0.66–0.95) that higher scores on grammar (CELF-4 Word Structure in Spanish), working memory (Memory Updating in Spanish), higher-level language ability (Inference Making Stories in Spanish), and English code-related skills (Letter Identification) were associated with a higher probability of correct responses. However, for each of these predictors, the posterior distributions also remain compatible with small, null, or negative effects. Figures 4 and 7 illustrate how model-derived effects translate into predicted probabilities of comprehension accuracy.

Robustness Checks

When multiple candidate measures were available for the same construct (e.g., PPVT-4 vs. CELF-4 Word Classes for vocabulary; Inference Making Stories vs. Picture Arrangement Task for higher-level language), we refit robustness check models substituting alternate measures to assess sensitivity to measurement choice (see Tables 3–4 and Supplemental Figures 3–4). Overall, effect estimates were similar across alternate measurement selections, indicating stability of the primary results. One notable pattern was that, for bilingual children, the predictive influence of Kindergarten vocabulary on Grade 6 English comprehension was stronger when vocabulary was measured with the English PPVT-4 than with the Spanish CELF-4 Word Classes measure — suggesting that predicting an English outcome benefits from including English vocabulary assessments.

We also checked whether, in analyzing Grade 6 reading comprehension for bilingual children, entering both Spanish and English measures would help to inform predictions of outcome comprehension accuracy. We did this only for the K predictors model because in PK we administered measures in Spanish, thus language comparisons were not possible. We found that the predictive effect of vocabulary or listening comprehension scores was much stronger for the English than for the Spanish measures of these constructs (see Supplemental Figure 5), even when measures in both languages were included in models. It is the case that English vocabulary and higher-level language measures in PK and K are most clearly associated with predicted change in reading comprehension outcome scores in English in Grade 6.

Because Bayesian inference depends on prior choices, we assessed prior sensitivity by refitting models under several reasonable prior specifications: Normal (0,1), Normal (0,10), and Student's t (3, 0, 10) for population- and group-level parameters, and LKJ(1) and LKJ(2) for correlation/covariance structures. Estimates and substantive conclusions remained stable across these specifications (see Supplemental Figures 6–7). Supplemental Figure 6 displays prior sensitivity results for the monolingual group and Supplemental Figure 7 for the bilingual group; Figure 8 provides an integrated visualization of these robustness checks.

Discussion

We identified prekindergarten and kindergarten language, cognitive, and code-related predictors associated with Grade 6 reading comprehension within monolingual and bilingual groups, characterizing group-specific patterns of association and informing targeted prevention strategies (LARRC, Yeomans-Maldonado, Bengochea, & Mesa, 2018).

For each group, early predictors showed a developmental shift from code-related skills in PK to broader language skills in kindergarten. Vocabulary emerged as the most consistent predictor, with additional contributions from grammar, listening comprehension, and, for bilingual children, cross-linguistic and cognitive factors.

Based on the literature, there are several reasons we may see different patterns of predictors between language groups. For vocabulary, bilingual children showed additional benefits from vocabulary in Spanish, which aligns with previous findings showing that bilingual students can benefit from cross-linguistic transfer to leverage their Spanish vocabulary to scaffold English reading comprehension (Grimm et al., 2018; LARRC et al., 2019; Proctor et al., 2006). Because bilingual children must process and integrate information across languages, it is not surprising that Spanish working memory, as assessed by memory updating, was a predictor of Grade 6 reading comprehension in the bilingual group. These group-specific patterns are consistent with theoretical and empirical work suggesting that reading comprehension develops through interacting linguistic and cognitive processes that may differ in timing or salience across language contexts.

The Centrality of Vocabulary

Our results illustrate the central role that vocabulary plays in subsequent reading comprehension beginning in the PK years. Similar results were reported by Grimm et al. (2017) in their longitudinal study of Spanish-speaking children and by Psyridou et al. (2021) and Manu et al. (2021) in their longitudinal studies of Finnish speaking children. In those studies, K vocabulary scores predicted significant variance in reading comprehension in 9th and 10th grades.

The importance of vocabulary for reading comprehension is well established - it was one of the “five pillars” essential to reading instruction identified by the National Reading Panel in 2000 (NICHD, NIH). What is concerning, given the centrality of vocabulary in academic achievement, is the lack of a consistent, systematic educational approach to increasing children’s vocabulary breadth and depth (Black & Wright, 2023). Challenges include identifying which words to teach (Wasik et al., 2016), how to teach them (Hadley & Mendez, 2021), a lack of clarity on effective instructional approaches (e.g., Cervetti et al., 2023; Dai et al., 2024) challenges with time availability in the school day (e.g. Wright & Neuman, 2014) and evidence that many bilingual children begin school knowing fewer words than their peers (e.g., Morgan et al., 2015; Rowe & Goldin-Meadow 2009).

Grammar and Morphology

In K, grammar and morphology were significant predictors of Grade 6 reading comprehension for the monolingual and bilingual groups, each in their primary language. In an earlier LARRC study examining the concurrent relationship between grammar and reading comprehension in the monolingual group, grammar was most closely related to reading comprehension in third grade (LARRC & Logan, 2017). Our results show that this relationship holds longitudinally at least through Grade 6.

Grammatical knowledge is thought to support a reader's ability to monitor syntactic information in the text for meaning and to generate meaning coherence during the reading process (Fender, 2001). Morphological awareness, including understanding inflectional and derivational morphemes, prefixes and suffixes, helps the reader decode multisyllabic words and understand their meaning (Levesque et al., 2019; Sparks & Metsala, 2023).

A recent meta-analysis of 86 studies investigating the relationship between grammatical knowledge and reading comprehension from elementary school through university in students who spoke Chinese, Dutch, English, and French showed a large effect size (Fisher's $z = 0.54$) with a significant interaction by grade (Zheng et al., 2023). The secondary school group had a significantly larger effect size than the primary grade group. These results are consistent with our finding that grammar and morphology in K predicted Grade 6 reading comprehension and are also consistent with longitudinal studies by Lervag et al. (2018), who reported that grammar assessed in Grade 2 Norwegian speaking children predicted reading comprehension in Grade 7 and Farnia and Geva (2013), who found that syntax in Grade 1 English language learners predicted English reading comprehension in Grade 6.

Just as systematic vocabulary instruction could benefit students who may be at risk for poor reading comprehension, systematic instruction for grammar and morphology can benefit young learners (Buggy & Dillon, 2025). For example, a 2010 systematic review on the effects of morphological instruction on literacy skills showed that this instruction was particularly beneficial for less able readers and was as effective for young students (e.g. PK, K) as it was for older students (Bowers et al, 2010). Morphological instruction was most effective when integrated with phonological, orthographic, grammatical, and semantic information.

Listening Comprehension

The importance of listening comprehension as a predictor of reading comprehension is emphasized in models such as the Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990). In reading research, listening comprehension is considered a higher-level language skill that integrates words, phrases, and sentences to form a mental model of the text. Research shows that the importance of listening comprehension, relative to code-related, increases over time as children become fluent readers (Garcia & Cain, 2014) and that vocabulary, inferencing, background knowledge, working memory, and attention all impact listening comprehension ability (Hogan et al., 2014). Kim and Pilcher (2016) suggested that "listening comprehension instruction should be an integral part of reading and writing instruction…" (p. 2) but also acknowledge that "…listening comprehension is a large problem space" (p.19). By this they mean that, like reading comprehension, listening comprehension is not a single skill that can be taught. It requires a systematic approach to building oral language skills that support oral language understanding embedded in language-rich environments. To this end, schools can evaluate whether their language arts, science, and math curricula promote oral language development at the discourse level (e.g., conversations, stories, discussions, debates) every day of the school year. Along with the lower level language skills of vocabulary, grammar, and morphology, this is crucial for skilled reading comprehension.

Letter Knowledge

In our study PK letter knowledge was an important predictor of Grade 6 reading comprehension for the monolingual group when all other predictors were in the model. Similar results were reported by Leppanen et al. (2008), Manu et al (2021), and Psyridou et al, (2021) in Finnish-speaking children. This builds on our work demonstrating the important role of letter knowledge for reading comprehension in Grade 3 via word reading (LARRC & Chiu, 2018). The indirect influence of letter knowledge on reading comprehension via word reading is supported by a systematic review of PK predictors of later reading comprehension which found that code-related skills, such as phoneme awareness and letter knowledge in PK, were indirectly related to reading comprehension via word reading (Hjetland et al., 2017). Together this research demonstrates that a strong foundation in code-related skills supports future word reading and reading comprehension.

Working Memory

Working memory plays an important role in reading comprehension by holding and updating the mental representation of information in the text referred to as the text base. Good comprehenders keep explicit information from the text and inferencing results activated to construct and update their own representation of text meaning, known as the situation model (Kintsch, 2009).

It is noteworthy that performance on a Spanish memory updating task by bilingual PK students was a significant predictor of reading comprehension seven years later in Grade 6. Previous longitudinal studies of reading comprehension have not typically assessed working memory and results for cross-sectional studies are mixed. Some have not found a direct relation between working memory and reading comprehension (e.g., LARRC & Logan, 2017; Oakhill & Cain, 2012) but some report a significant relationship. For example, Linares and Pelegrina (2023) assessed fourth graders in Spain and found that after controlling for fluid intelligence, reading and math fluency, vocabulary and reading ability, the retrieval component of working memory updating was directly related to reading comprehension. A recent meta-analysis by Shin (2020) examined the relationship between performance on variations of the reading span task commonly used to assess working memory where participants listen to or read a series of sentences and make some decision about the sentences such as whether each is grammatically correct, while holding the final word of the sentence in memory. Shin reported a moderate relationship between reading comprehension in L2 and working memory (r = .30).

Limitations

Several limitations should be considered when interpreting these findings. First, the monolingual and bilingual samples differed in their geographic recruitment. Monolingual English-speaking children were recruited from multiple states, whereas the Spanish–English bilingual group was recruited solely from Arizona. This difference may limit the generalizability of findings for bilingual learners to other Spanish-speaking populations in the United States. At the same time, recruiting the bilingual sample from a single state resulted in a relatively cohesive linguistic group, with most children speaking Mexican Spanish and experiencing similar educational language policies. This reduced dialectal and sociolinguistic variability may have increased internal validity for analyses conducted within the bilingual group. Importantly, because prior work with these samples has demonstrated that oral language and reading comprehension constructs are not measurement-invariant across bilingual and monolingual children, we did not conduct direct cross-group comparisons. Accordingly, differences in geographic recruitment should not be interpreted as a confound affecting between-group contrasts but rather as contextual factors relevant to understanding group-specific developmental patterns.

Educational Implications

The present findings do not provide evidence for the effectiveness of specific instructional approaches; however, they inform the selection of early screening measures that may help identify children who could benefit from additional instructional support. Based on our results, we recommend different assessment approaches by population and grade. For children who enter PK speaking monolingual English, we suggest the WRMT-R/NU Letter ID subtest; for children who enter PK speaking primarily Spanish, we suggest the PPVT-4 or the CELF-4 Spanish Understanding Spoken Paragraphs subtest. In kindergarten, we suggest the PPVT-4 or the CELF-4 Word Structure subtest for monolingual English-speaking children, and the PPVT-4 and the CELF-4 Spanish Word Structure subtest for Spanish–English bilingual children. The predictive utility of these measures may vary across samples, outcome measures, and test editions; therefore, they should be used as screening tools to guide instruction rather than for diagnostic classification.

These assessments are often used in psychoeducational evaluations, but they may also serve as screening tools to identify children who could benefit from additional instructional support. Importantly, this use is distinct from identifying risk for a language or literacy disorder. Measures such as the PPVT may yield lower scores in linguistically and culturally diverse populations (e.g., Restrepo et al., 2006; Washington & Craig, 1992) and do not accurately identify language impairment (Gray et al., 1999). Accordingly, these tools should be interpreted within culturally and linguistically responsive frameworks. When used appropriately, they can support educators in identifying areas for targeted instruction, including vocabulary, grammar, and code-related skills, and in promoting equitable literacy outcomes.

Artificial Intelligence Statement: ChatGPT was used during the revision process for this manuscript to reduce the number of words in the manuscript and to review the manuscript for clarity, redundancy, and formatting. The resulting text was reviewed and revised by the authors prior to inclusion in the manuscript

Acknowledgements

LARRC work was supported by Grant R305F100002 of the Institute of Education Sciences' Reading for Understanding Initiative. MBRC work was supported by National Institutes of Health NICHD Grant R. We are deeply grateful to the numerous staffs, research associates, school administrators, teachers, children, and families who participated. The views presented in this work do not represent those of the federal government, nor do they endorse any products or findings presented in this article.

Declaration of Interest Statement:

The authors report there are no competing interests to declare.

Supplemental Material

Supplemental material is available to accompany this manuscript.

Data Availability Statement

Portions of the data from this study may be available upon reasonable request from the study authors.

References

Arıcı, B., & Faculty, E. (2020). Efficiency of reading time application in schools through the eyes of students. *Educational Policy Analysis and Strategic Research*. https://doi.org/10.29329/epasr.2020.323.7.

August, D., Artzi, L., Barr, C., & Francis, D. (2018). The moderating influence of instructional intensity and word type on the acquisition of academic vocabulary in young English language learners. *Reading and Writing, 31*, 965-989. https://doi.org/10.1007/S11145-018-9821-1.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Belacchi, C., Carretti, B., & Cornoldi, C. (2010). The role of working memory and updating in Coloured Raven Matrices performance in typically developing children. *European Journal of Cognitive Psychology, 7*, 1010-1020. https://doi.org/10.1080/09541440903184617

Bell, N., & Wheldall, K. (2022). Factors contributing to reading comprehension in children with varying degrees of word-level proficiency. *Australian Journal of Education, 66*, 73-91. https://doi.org/10.1177/00049441211062941

Berry, E. D. J., Allen, R. J., Mon-Williams, M., & Waterman, A. H. (2019). Cognitive offloading: structuring the environment to improve children's working memory task performance. *Cognitive Science, 43*, e12770. https://doi.org/10.1111/cogs.12770

Black, C., & Wright, K. (2023). What's up with words? A systematic review of designs, strategies, and theories underlying vocabulary research. *Reading Psychology, 45*, 78 - 103. https://doi.org/10.1080/02702711.2023.2253249.

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research, 80*, 144-179. DOI: 10.3102/0034654309359353

Buggy, R. & Dillon, B. (2025). The science of syntax and predictability. *Perspectives on Language and Literacy, 51*, 12-15. http://www.onlinedigeditions.com/publication/?i=847535&p=1&view=issueViewer

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing: An Interdisciplinary Journal, 11*, 489–503. https://doi.org/10.1023/A:1008084120205

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*, 31-42. https://doi.org/10.1037/0022-0663.96.1.31

Cain, K., Currie, N. K., Francey, G., Davies, R., Gray, S., Bridges, M. S., Restrepo, M. A., Thompson, M. S., & Ciraolo, M. F. (in press). The influence of reader and text characteristics on sixth graders' inference making. *Journal of Research in Reading. https://doi.org/10.1111/1467-9817.12474*

Cajkler, W. & Hislam, J. (2002). Trainee teachers' grammatical knowledge: The tension between public expectations and individual competence. *Language Awareness, 11*, 161–177. https://doi.org/10.1080/09658410208667054

Cajkler, W., & Hislam, J. (2004). Teaching and understanding grammar: How trainee teachers make sense of a new curriculum? In N. Bartels (Ed.), *Applied Linguistics in Language Teacher Education*, 295–312. Dordrecht: Kluwer.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest, 19*, 5–51. https://doi.org/10.1177/1529100618772271

Cervetti, G., Fitzgerald, M., Hiebert, E., & Hebert, M. (2023). Meta-analysis examining the impact of vocabulary instruction on vocabulary knowledge and skill. *Reading Psychology, 44*, 672 - 709. https://doi.org/10.1080/02702711.2023.2179146.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioural sciences (3rd. Edition). Lawrence Erlbaum Associates.

Colmar, S., Davis, N. & Sheldon, L. (2016). A pilot classroom-based study of attention and working memory strategies for primary-aged students. *Journal of Psychologists and Counsellors in Schools, 26*, 125-136.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*, 222–251. https://doi.org/10.3102/00346543049002222

Currie, N. K., & Cain, K. (2015). Children's inference generation: The role of vocabulary and working memory. *Journal of Experimental Child Psychology, 137*, 57–75. https://doi.org/10.1016/j.jecp.2015.03.005

Currie, N. K., Francey, G., Davies, R., Gray, S. Bridges, M. S., Restrepo, M. A., Thompson, M. S., Ciralo, M. F., Hu, J., & Cain, K. (2021). The process and product of coherence monitoring in young readers: Effects of reader and text characteristics. *Scientific Studies of Reading, 25*, 141-158. doi: 10.1080/10888438.2020.1831503

Dai, Y., Su, Y., & Li, X. (2024). Exploring teaching methods for middle school English vocabulary in China: Based on observation of high-quality lessons in people's education edition. *Global Journal of Social Sciences Studies*. https://doi.org/10.55284/gjss.v10i1.1136.

de Jong, P. F., & Smit, L. A. (2020). A beneficial effect of a Cogmed working memory training on intelligence without an improvement in working memory capacity. *Psychology & Neuroscience, 13*, 341–356. https://doi.org/10.1037/pne0000196

Dunn, L.M. & Dunn, D.M. (2007). PPVT-4: Peabody picture vocabulary test: Pearson assessments.

Eager, C., & Roy, J. (2017). Mixed effects models are sometimes terrible. arXiv preprint arXiv:1701.04858.

Farnia, F. & Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading, 36*, 389–421. https://doi.org/10.1111/jrir.12003

Fender M. (2001). A review of L1 and L2/ESL word integration skills and the nature of L2/ESL word integration development involved in lower-level text processing. *Language Learning, 51*, 319–396. doi: 10.1111/0023-8333.00157

Friesen, D., Schmidt, K., Atwal, T., & Celebre, A. (2022). Reading comprehension and strategy use: Comparing bilingual children to their monolingual peers and to bilingual adults. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.986937.

Fuchs, D., Compton, D. L., Fuchs, L. S., Bryant, V. J., Hamlett, C. L., & Lambert, W. (2012). First-grade cognitive abilities as long-term predictors of reading comprehension and disability status. *Journal of Learning Disabilities, 45*, 217-231. https://doi.org/10.1177/0022219412442154 (Original work published 2012)

García, J. R. and Cain, K. (2014). Code-related and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, *84*, 74-111. https://doi.org/10.3102/0034654313499616

Gough, P. B., & Tunmer,W. E. (1986). Code-related, reading, and reading disability. *Remedial and Special Education, 7*, 6–10. doi:10.1177/074193258600700104

Gray, S., Plante, E., Vance, R., & Henrichsen, M. (1999). The diagnostic accuracy of four vocabulary tests administered to Prekindergarten-age children. *Language, Speech, and Hearing Services in Schools, 30*, 196-206. https://doi.org/10.1044/0161-1461.3002.196.

Grimm, R. P., Solari, E. J., & Gerber, M. M. (2018). A longitudinal investigation of reading development from kindergarten to grade eight in a Spanish-speaking bilingual population. *Reading & Writing, 31*, 559–581. https://doi.org/10.1007/s11145-017-9798-1

Hadley, E. B. & Mendez, Z., (2021). A systematic review of word selection in early childhood vocabulary instruction. *Early Childhood Research Quarterly, 54,* 44-59. https://doi.org/10.1016/j.ecresq.2020.07.010

Hjetland H. N., Brinchmann E. I., Scherer R., & Melby-Lervåg M. (2017). Prekindergarten predictors of later reading comprehension ability: a systematic review. Campbell Systematic Reviews 2017:14. DOI: 10.4073/csr.2017.14

Hogan, T., Adlof, S. A., & Alonzo, C. (2014). On the importance of listening comprehension. *International Journal of Speech-Language Pathology, 16,* 199-207. 10.3109/17549507.2014.904441

Holmes J, Gathercole SE, Dunning DL. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science; 12*, F9–F15. https://doi.org/10.1111/j.1467-7687.2009.00848.x

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2*, 127–160. doi:10.1007/bf00401799

Humairoh, S. (2022). Efforts to increase reading interest in elementary schools. Proceedings of the 1st International Conference on Social Science (ICSS). https://doi.org/10.59188/icss.v1i1.58.

Jones, J. S., Milton, F., Mostazir, M., & Adlam, A. R. (2020). The academic outcomes of working memory and metacognitive strategy training in children: A double-blind randomized controlled trial. *Developmental Science, 23*, 17. https://doi.org/10.1111/desc.12870

Jones, S., Myhill, D. & Bailey, T. (2013). Grammar for writing? An investigation of the effects of contextualized grammar teaching on students' writing. *Reading and Writing, 26,* 1241-1263. DOI 10.1007/s11145-012-9416-1

Joshi, R. M., Ji, X. R., Breznitz, Z., Amiel, M., & Yulia, A. (2015). Validation of the Simple View of Reading in Hebrew—A Semitic Language. *Scientific Studies of Reading, 19*, 243–252. https://doi.org/10.1080/10888438.2015.1010117

Kaufman, A.S., & Kaufman, N.L. (1997). Kaufman Brief Intelligence Test - Second edition, Minneapolis, MN: NCS Pearson.

Kelley, J., Lesaux, N., Kieffer, M., & Faller, E. (2010). Effective academic vocabulary instruction in the urban middle school. *The Reading Teacher, 64*, 5-14. https://doi.org/10.1598/RT.64.1.1.

Kershaw, S., & Schatschneider, C. (2012). A latent variable approach to the simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 433–464. https://doi.org/10.1007/s11145-010-9278-3

Kim, Y. S. G. & Pilcher, H. (2016). What is listening comprehension and what does it take to improve listening comprehension? In R. Schiff & M. Joshi (Eds.), Handbook of interventions in learning disabilities (pp. 159-174). New York: Springer.

Kintsch, W. (2009). Learning and constructivism. In S. Tobias & T. M. Duffy (Eds.), Constructivist Instruction: Success or failure? New York, NY: Routledge.

Kulesz, P. A., Francis, D. J., Barnes, M. A., & Fletcher, J. M. (2016). The influence of properties of the test and their interactions with reader characteristics on reading comprehension: An explanatory item response study. *Journal of Educational Psychology, 108*, 1078–1097. https://doi.org/10.1037/edu0000126

Language and Reading Research Consortium (2015a). The dimensionality of Spanish in young Spanish-English dual language learners. *Journal of Speech, Language, and Hearing Research, 58*, 754-766. DOI: 10.1044/2015_JSLHR-L-13-0266.

Language and Reading Research Consortium (2015b). The dimensionality of language ability in young children. *Child Development, 86*, 1948-1965. DOI: http://dx.doi.org/10.1111/cdev.12450.

Language and Reading Research Consortium (2015c). Learning to read: should we keep things simple? *Reading Research Quarterly, 50*, 151-169. DOI: 10.1002/rrq.99

Language and Reading Research Consortium, Pratt, A., & Logan, J. (2014). Improving language-focused comprehension instruction in primary grade classrooms: Impacts of the Let's Know! Experimental curriculum. *Educational Psychology Review, 26*, 357–377. https://doi.org/10.1007/s10648-014-9275-1

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. https://doi.org/10.1044/2017_JSLHR-L-16-0039

Language and Reading Research Consortium & Logan, J. (2017). Pressure points in reading comprehension: A quantile multiple regression analysis. *Journal of Educational Psychology, 109*, 451-464. http://dx.doi.org/10.1037/edu0000150

Language and Reading Research Consortium (LARRC) & Yeomans-Maldonado, G. (2017). Development of comprehension monitoring in beginner readers. *Reading and Writing: An Interdisciplinary Journal, 30*, 2039-2067. doi:10.1007/s11145-017-9765-x

Language and Reading Research Consortium & McIlraith, A. L. (2018). Predicting word reading ability: A quantile regression study. *Journal of Research in Reading, 41*, 79-96. https://doi.org/10.1111/1467-9817.12089

Language and Reading Research Consortium (LARRC), Jiang, H., Logan, J. A., & Jia, R. (2018). Modeling the nature of grammar and vocabulary trajectories from pre-kindergarten to third grade. *Journal of Speech, Language, and Hearing Research, 61*, 910-923. doi: 10.1044/2018_JSLHR-L-17-0090. PubMed PMID: 29642241.

Language and Reading Research Consortium (LARRC), Yeomans-Maldonado, G., Bengochea, A., & Mesa, C. (2018). The dimensionality of oral language in kindergarten Spanish-English dual-language learners. *Journal of Speech, Language, and Hearing Research, 61*, 2779-2795. https://doi.org/10.1044/2018_jslhr-l-17-0320

Language and Reading Research Consortium, Farquharson, K., & Jiang, H. (2018). Are working memory and attention equally important for reading and listening comprehension? A developmental comparison. *Reading and Writing, 31*, 1449-1477. https://doi.org/10.1007/s11145-018-9840-y

Language and Reading Research Consortium (LARRC) & Chiu, Y. D. (2018). The simple view of reading across development: The prediction of grade 3 reading comprehension by prekindergarten skills. *Remedial and Special Education, 39*, 289-303. https://doi.org/10.1177/0741932518762055

Language and Reading Research Consortium (LARRC) & Muijselaar, M. M. L. (2018). The dimensionality of inference making: Are local and global inferences distinguishable? *Scientific Studies of Reading, 22*, 117-136. https://doi.org/10.1080/10888438.2017.1371179

Language and Reading Research Consortium (LARRC), Currie, N. K., & Muijselaar, M. M. L. (2019). Inference making in young children: The concurrent and longitudinal contributions of verbal working memory and vocabulary. *Journal of Educational Psychology, 111*, 1416-1431. https://doi.org/10.1037/edu0000342

Language and Reading Research Consortium (LARRC), Mesa, & Yeoman's-Maldonado (2019). The role of prek Spanish in predicting first grade English word reading among dual language learners. *Journal of Speech, Language, and Hearing Research, 62*, 1755-1774. https://doi.org/10.1044/2019_JSLHR-L-17-0146

Language and Reading Research Consortium (LARRC), Mesa, C., & Yeomans-Maldonado, G. (2021). English and Spanish predictors of grade 3 reading comprehension in bilingual children. *Journal of Speech, Language, and Hearing Research, 64*, 889-908. https://doi.org/10.1044/2020_jslhr-20-00379

Lervåg, A., Dolean, D., Tincas, I., & Melby-Lervåg, M. (2019). Socioeconomic background, nonverbal IQ and school absence affects the development of vocabulary and reading comprehension in children living in severe poverty. *Developmental Science*, 22, e12858.

Lervåg, A., Hulme, C., & Melby-Lervåg, M. (2018). Unpicking the developmental relationship between oral language skills and reading comprehension: It's simple, but complex. *Child Development, 89*, 1821-1838.

Levesque, K. C., Kieffer, M. J., and Deacon, S. H. (2019). Inferring meaning from meaningful parts: The contributions of morphological skills to the development of children's reading comprehension. *Reading Research Quarterly, 54*, 63–80. doi: 10.1002/rrq.219

Liddell, T. M., & Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong?. *Journal of Experimental Social Psychology, 79*, 328-348.

Linan-Thompson, S., Vaughn, S., Prater, K., & Cirino, P. T. (2006). The response to intervention of English language learners at risk for reading problems. *Journal of Learning Disabilities, 39*, 390-398. https://doi.org/10.1177/00222194060390050201

Linares, R. & Pelegrina, S. (2023). The relationship between working memory updating components and reading comprehension. *Cognitive Processing, 24*, 253-265. https://doi.org/10.1007/s10339-023-01127-3

Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority learners. *Journal of Educational Psychology, 102*, 701–711. https://doi.org/10.1037/a0019135

Manu, M., Torppa, M., Eklund, K., Poikkeus, A., Lerkkanen, M, & Niemi, P. (2021). Kindergarten pre-reading skills predict Grade 9 reading comprehension (PISA Reading) but fail to explain gender difference. *Reading and Writing: An Interdisciplinary Journal, 34*, 753–771. https://doi.org/10.1007/s11145-020-10090-w

Marulis, L., & Neuman, S. (2010). The effects of vocabulary intervention on young children's word learning. *Review of Educational Research, 80*, 300 - 335. https://doi.org/10.3102/0034654310377087.

McCardle, P., Scarborough, H. S., & Catts, H. W. (2001). Predicting, explaining, and preventing children's reading difficulties. *Learning Disabilities Research & Practice, 16*, 230-239. https://doi.org/10.1111/0938-8982.00023

McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children, 71*, 445-463. https://doi.org/10.1177/001440290507100404

McMurtrie, B. (2024, May 9). Is this the end of reading? The Chronicle of Higher Education. https://www.chronicle.com/article/is-this-the-end-of-reading

McQuillan, J. L. (2019). The inefficiency of vocabulary instruction. International Electronic *Journal of Elementary Education, 11*, 309-318. DOI: 10.26822/iejee.2019450789

Miller, A. C., Davis, N., Gilbert, J. K., Cho, S.-J., Toste, J. R., Street, J., & Cutting, L. E. (2014). Novel Approaches to Examine Passage, Student, and Question Effects on Reading Comprehension. *Learning Disabilities Research & Practice*: A Publication of the Division for Learning Disabilities, Council for Exceptional Children, 29(1), 25–35. https://doi.org/10.1111/ldrp.12027

Moats, L. (2009). Knowledge foundations for teaching reading and spelling. *Reading and Writing, 22*, 379-399. DOI 10.1007/s11145-009-9162-1

Morgan, P., Farkas, G., Hillemeier, M., Hammer, C., & Maczuga, S. (2015). 24-Month-Old Children With Larger Oral Vocabularies Display Greater Academic and Behavioral Functioning at Kindergarten Entry. *Child development, 86*, 1351-70. https://doi.org/10.1111/cdev.12398.

Muñez, D., Lee, K., Bull, R., Khng, K. H., Cheam, F., & Rahim, R. A. (2022). Working memory and numeracy training for children with math learning difficulties: Evidence from a large-scale implementation in the classroom. *Journal of Educational Psychology, 114*, 1866.

Muñoz-Sandoval, A. F., Woodcock, R. W., McGrew, K. S., & Mather, N. (2005). Batería III Woodcock-Muñoz: Pruebas de habilidades cognitivas. Itasca, IL: Riverside Publishing.

Murphy, K., Language and Reading Research Consortium, Farquharson, K. (2016). Investigating profiles of lexical quality in Prekindergarten and their contribution to first grade reading. *Reading and Writing: An Interdisciplinary Journal, 29*, 1745-1770. https://doi.org/10.1007/s11145-016-9651-y

Nation, K. (2019). Children’s reading difficulties, language, and reflections on the simple view of reading. *Australian Journal of Learning Difficulties, 24*, 47 - 73. https://doi.org/10.1080/19404158.2019.1609272.

National Reading Panel. (2000). Report of the National Reading Panel: Teaching children to read. Reports of the subgroups (NIH Publication No. 00-4754). National Institute of Child Health and Human Development.

Nalborczyk, L., Batailler, C., Lœvenbruck, H., Vilain, A., & Bürkner, P. C. (2019). An Introduction to Bayesian Multilevel Models Using brms: A Case Study of Gender Effects on Vowel Variability in Standard Indonesian. *Journal of Speech, Language, and Hearing Research, 62*, 1225–1242. https://doi.org/10.1044/2018_JSLHR-S-18-0006

Oakhill, J., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language, Cognition and Neuroscience, 18,* 443-468. DOI:10.1080/01690960344000008

Oakhill, J. V. & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading, 16*, 91–121. https://doi.org/10.1080/10888438.2010.529219

Ortiz, M., Folsom, J. S., Al Otaiba, S., Greulich, L., Thomas-Tate, S., & Connor, C. M. (2012). The componential model of reading: Predicting first grade reading performance of culturally diverse students from ecological, psychological, and cognitive factors assessed a kindergarten entry. *Journal of Learning Disabilities, 45*, 406-417. doi: 10.1177/0022219411431242

Paterson, L. L. (2010). Grammar and the English National Curriculum. *Language and Education, 24*, 473-484. https://doi.org/10.1080/09500782.2010.495782

Peng, P., Fuchs, D., Fuchs, L. S., Elleman, A. M., Kearns, D. M., Gilbert, J. K., Compton, D. L., Cho, E., & Patton, S. (2019). A longitudinal analysis of the trajectories and predictors of word reading and reading comprehension development among at-risk readers. *Journal of Learning Disabilities*, 52, 195-208. https://doi.org/10.1177/0022219418809080

Peters-Sanders, L., Sanders, H., Goldstein, H., & Ramachandran, K. (2023). using multivariate adaptive regression splines to predict lexical characteristics' influence on word learning in first through third graders. *Journal of Speech, Language, and Hearing Research*, 1-16. https://doi.org/10.1044/2022_JSLHR-22-00165.

Proctor, C. P., August, D., Carlo, M. S., & Snow, C. (2006). The intriguing role of Spanish language vocabulary knowledge in predicting English reading comprehension. *Journal of Educational Psychology, 98*, 159–169. https://doi.org/10.1037/0022-0663.98.1.159

Proctor, C., Silverman, R., Harring, J., & Montecillo, C. (2012). The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish–English bilingual children in elementary school. *Reading and Writing, 25*, 1635-1664. https://doi.org/10.1007/S11145-011-9336-5.

Protopapas, A., Simos, P., Sideridis, G. D., & Mouzaki, A. (2012). The components of the Simple View of Reading: A confirmatory factor analysis. *Reading Psychology, 33*, 217-240. http://dx.doi.org/10.1080/02702711.2010.507626

Psyridou, M., Tolvanen, A., Jong, P., Lerkkanen, M., Poikkeus, A. & Torppa, M. (2021). Developmental profiles of reading fluency and reading comprehension from grades 1 to 9 and their early identification. *Developmental Psychology. 57*, 1840-1854. 10.1037/dev0000976.

Pikulski, J. J. (1994). Preventing reading failure: A review of five effective programs. *Reading Teacher, 48*, 30-39.

Restrepo, M., Schwanenflugel, P., Blake, J., Neuharth-Pritchett, S., Cramer, S., & Ruston, H. (2006). Performance on the PPVT-III and the EVT: Applicability of the measures with African American and European American Prekindergarten children. *Language, Speech, and Hearing Services in Schools, 37*, 17-27. https://doi.org/10.1044/0161-1461(2006/003).

Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin & Review*, *26*(2), 452–467. https://doi.org/10.3758/s13423-018-1558-y

Rowe, M., & Goldin-Meadow, S. (2009). Differences in Early Gesture Explain SES Disparities in Child Vocabulary Size at School Entry. *Science, 323*, 951 - 953. https://doi.org/10.1126/science.1167025.

Sala, G., & Gobet, F. (2017). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current Directions in Psychological Science, 26*, 515-520.

Sánchez-Pérez, N., Castillo, A., López-López, J. A., Pina, V., Puga, J. L., Campoy, G., ... & Fuentes, L. J. (2018). Computer-based training in math and working memory improves cognitive skills and academic achievement in primary school children: Behavioral results. *Frontiers in Psychology, 8*, 2327.

Schrank, F. A. (2006). Specification of the cognitive processes involved in performance on the Woodcock-Johnson III (Assessment Service Bulletin No. 7). Itasca, IL: Riverside Publishing.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). Clinical evaluation of language fundamentals: Examiner's manual (4th ed.). Bloomington, MN: Pearson, Inc.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). Clinical Evaluation of Language Fundamentals-Spanish Edition; Examiner's Manual. Bloomington, MN: Pearson, Inc.

Shields, K. A., Cook, K. D., & Greller, S. (2016). How kindergarten entry assessments are used in public schools and how they correlate with spring assessments (REL 2017–182). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast Islands. http://ies.ed.gov/ncee/edlabs/projects/project.

Shin, J. (2020). A meta-analysis of the relationship between working memory and second language reading comprehension: Does task type matter? *Applied Psycholinguistics, 41,* 873-900. *doi:10.1017/S0142716420000272*

Silva, M., & Cain, K. (2015). The relations between lower and higher level comprehension skills and their role in prediction of early reading comprehension. Journal of Educational Psychology, 107, 321–331. https://doi.org/10.1037/a0037769

Silverman, R., Proctor, C., Harring, J., Hartranft, A., Doyle, B., & Zelinke, S. (2015). Language skills and reading comprehension in English monolingual and Spanish–English bilingual children in grades 2–5. *Reading and Writing, 28*, 1381-1405. https://doi.org/10.1007/S11145-015-9575-Y.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.) (1998). Preventing reading difficulties in young children. Washington, DC: National Academies Press.

Snow, C. E., & Juel, C. (2005). Teaching children to read: What do we know about how to do it? In M. J. Snowling & C. Hulme (Eds.), *The Science of Reading*. Blackwell.

Sparks, E., & Metsala, J. L., (2023). Morphological awareness predicts reading comprehension in first grade students. *Frontiers in Education, 8.* *https://doi.org/10.3389/feduc.2023.1304688*

Sparks, R. L., Patton, J. & Murdoch, A. (2014). Early reading success and its relationship to reading achievement and reading volume: Replication of '10 years later.' *Reading and Writing, 27*, 189–211. https://doi.org/10.1007/s11145-013-9439-2

Spencer, M., Gilmour, A. F., Miller, A. C., Emerson, A. M., Saha, N. M., & Cutting, L. E. (2019). Understanding the influence of text complexity and question type on reading outcomes. *Reading and Writing, 32* 603–637. https://doi.org/10.1007/s11145-018-9883-0

Spencer-Smith, M., Quach, J., Mensah, F., Roberts, G., Gathercole, S., Wake, M., & Anderson, P. J. (2020). The effectiveness of working memory training for children with low working memory. *Pediatrics, 146*, e20194028. https://doi.org/10.1542/peds.2019-4028

Tegmark, M., Alatalo, T., Vinterek, M., & Winberg, M. (2022). What motivates students to read at school? Student views on reading practices in middle and lower-secondary school. *Journal of Research in Reading*. https://doi.org/10.1111/1467-9817.12386.

Thorell, L. B., Lindqvist, S., Nutley, S. B., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in Prekindergarten children. *Developmental Science, 12,* 106–113. https://doi.org/10.1111/j.1467-7687.2008.00745.x

Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., & Rapp, D (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading, 32,* 383-401. https://doi.org/10.1111/j.1467-9817.2009.01401.x

Vernucci, S., Canet Juric, L., Introzzi, I., & Richard's, M. M. (2020). Working memory training in children: A review of basic methodological criteria. *Psychological Reports, 123,* 605-632. https://doi.org/10.1177/0033294119832978

Washington, J., & Craig, H. (1992). Performances of low-income, african american Prekindergarten and kindergarten children on the peabody picture vocabulary test-revised. *Language Speech and Hearing Services in Schools, 23*, 329-333. https://doi.org/10.1044/0161-1461.2304.329.

Wasik, B., Hindman, A., & Snell, E. (2016). Book reading and vocabulary development: A systematic review. *Early Childhood Research Quarterly, 37*, 39-57. https://doi.org/10.1016/J.ECRESQ.2016.04.003.

Wechsler, D. (1992). Wechsler intelligence scale for children (3rd ed.). New York, NY: The Psychological Corporation.

Woodcock, R. W. (1998). Woodcock Reading Mastery Tests – Revised/Normative Update. Circle Pines, MN: American Guidance Service/Pearson Assessments.

Wright, T., & Neuman, S. (2014). Paucity and disparity in kindergarten oral vocabulary instruction. *Journal of Literacy Research, 46*, 330 - 357. https://doi.org/10.1177/1086296X14551474.

Zhang, P. (2022). How does repetition affect vocabulary learning through listening to the teacher's explicit instruction? The moderating role of listening proficiency and preexisting vocabulary knowledge. *Language Teaching Research*. https://doi.org/10.1177/13621688221140521.

Zheng, H., Miao, X., Dong, Y., & Yuan, D. C. (2023). The relationship between grammatical knowledge and reading comprehension: A meta-analysis. *Frontiers in Psychology*, *14*. doi:10.3389/fpsyg.2023.1098568

Figure Captions

Figure 1

Mean Reading Comprehension Accuracy on the WRMT-R/NU Update Passage Comprehension Subtest by Child and by Test Item for the Monolingual Group

Note. Each point on the plot represents the mean response accuracy, by child or by item, in the reading comprehension measure. Means are calculated over all responses made by a child (for all items, to get the mean by child) or over all responses made to an item (made by all children, to get the mean by item). We have jittered the position of the points (adding a little random noise to the coordinates) because there may be multiple (by-child, or by-item) scores for some mean accuracy values.

Figure 2

Mean Reading Comprehension Accuracy on the WRMT-R/NU Update Passage Comprehension Subtest by Child and by Test Item for the Bilingual Group

Note. Each point on the plot represents the mean response accuracy, by child or by item, in the reading comprehension measure. Means are calculated over all responses made by a child (for all items, to get the mean by child) or over all responses made to an item (made by all children, to get the mean by item). We have jittered the position of the points (adding a little random noise to the coordinates) because there may be multiple (by-child, or by-item) scores for some mean accuracy values.

Figure 3
Posterior Probability Distribution for Prekindergarten Predictors of Reading Comprehension in the Monolingual Group

Note: Dots represent the posterior median. Lines represent the 66% (thicker line) and 95% (thinner line) credible intervals (Crl).

Figure 4
Posterior Probability Distribution for Kindergarten Predictors of Reading Comprehension in the Monolingual Group

Note. Dots represent the posterior median. Lines represent the 66% (thicker line) and 95% (thinner line) credible intervals.

Figure 5
Model Predictions for the Monolingual Group

Note. The plots represent the central (median) predicted change in the probability of outcome response accuracy, together with the 95% (light shaded ellipse) and the 66% (darker shaded ellipse) credible intervals. Plots indicate the predicted probability of a correct response given variation in code-related test scores (panel 1); and for the kindergarten predictors model, given variation in grammar (panel 2), vocabulary (panel 3) and listening comprehension test scores (panel 4).

Figure 6
Posterior Probability Distribution for Prekindergarten Predictors of Reading Comprehension in the Bilingual Group

Note. Dots represent the posterior median. Lines represent the 66% (thicker line) and 95% (thinner line) credible intervals (Crl).

Figure 7
Posterior Probability Distribution for Kindergarten Predictors of Reading Comprehension in the Bilingual Group

Note. Dots represent the posterior median. Lines represent the 66% (thicker line) and 95% (thinner line) credible intervals.

Figure 8
Model Predictions for the Bilingual Group

Note. The plots represent the central (median) predicted change in the probability of outcome response accuracy, together with the 95% (light shaded ellipse) and the 66% (darker shaded ellipse) credible intervals. Plots indicate the predicted probability of a correct response for the Prekindergarten predictors model, given variation in code-related test scores (panel 1); and for the kindergarten predictors model, given variation in grammar (panel 2), vocabulary (panel 3) and listening comprehension test scores (panel 4).

Supplemental Figure 1
Predictor Variable Density Plots for the Monolingual Group at Each Measurement Time Point

Note. Scores range from 0 to the sample maximum for each variable. The height of the density curve reflects the number of children with a given score. The one-dimensional rug plot beneath each curve indicates observed values, with darker ticks representing multiple children sharing the same score. Plot titles indicate the construct and corresponding measurement.

Supplemental Figure 2
Predictor Variable Density Plots for the Bilingual Group at Each Measurement Time Point

Note. Scores range from 0 to the sample maximum for each variable. The height of the density curve reflects the number of children with a given score. The one-dimensional rug plot beneath each curve indicates observed values, with darker ticks representing multiple children sharing the same score. Plot titles indicate the construct and corresponding measurement.

Supplemental Figure 3
Posterior Probability Distributions for Prekindergarten and Kindergarten Models Including Different Measures of Vocabulary and Higher-Level Language Constructs

Note. Light green represents estimates from the primary model (reported in the main text), and dark green represents estimates from alternate models including different measures of vocabulary and higher-level language. Dots indicate posterior medians, and lines represent 95% credible intervals (CrI).

Supplemental Figure 4
Posterior Probability Distributions for Prekindergarten and Kindergarten Models Including Alternate Measures of Vocabulary and Higher-Level Language Constructs

Note. Light blue represents estimates from the primary model (reported in the main text), and dark blue represents estimates from alternate models. Dots represent the posterior median. Lines represent the 95% credible intervals (Crl).

Supplemental Figure 5
Posterior Probability Distributions for Kindergarten Models Including Both English and Spanish Measures

Note. In light blue we show the primary model estimates (reported in the main text) versus in dark blue alternate model estimates, for models incorporating both English and Spanish measures of vocabulary and listening comprehension. Dots represent the posterior median. Lines represent the 95% quantile credible intervals.

Supplemental Figure 6
Posterior Probability Distributions for Models with Prekindergarten or Kindergarten Measures as Predictors for the Monolingual Group

Note. Different shades of green represent estimates from models fitted under alternative prior specifications. Dots indicate posterior medians, and lines represent 95% credible intervals (CrI).

Supplemental Figure 7
Posterior Probability Distributions for Models with Prekindergarten or Kindergarten Measures as Predictors for the Bilingual Group

Note. Different shades of blue represent estimates from models fitted under different alternative prior specifications. Dots represent the posterior median. Lines represent the 95% credible intervals (Crl).