HHS Public Access
Author manuscript
J Educ Psychol. Author manuscript; available in PMC 2021 April 01.

Published in final edited form as:
J Educ Psychol. 2020 April ; 112(3): 608–627. doi:10.1037/edu0000382.

Differential Co-Development of Vocabulary Knowledge and Reading Comprehension for Students with and without Learning Disabilities

Jamie M. Quinn,
Florida State University

Richard K. Wagner,
Florida State University

Yaacov Petscher,
Florida State University

Greg Roberts,
The University of Texas at Austin

Andrew J. Menzel,
RiskRecon, LLC

Christopher Schatschneider
Florida State University

Abstract

In this large-scale study of students from Title 1 schools ($N = 14{,}773$), we used multiple-group latent change score (LCS) modeling to investigate the developmental relations between vocabulary knowledge and reading comprehension in students with a school-identified learning disability (LD, $n = 627$) and typically developing students ($n = 14{,}146$). Students were tested for their vocabulary breadth and passage comprehension skills in Kindergarten through fourth grade. For typically developing students, there were bidirectional influences between their vocabulary knowledge and reading comprehension skills. There were no cross-lagged influences across constructs for students with an LD. We find evidence for a developmental delay, such that students with an LD had similar levels and gains in their vocabulary knowledge relative to typically developing students, but these students started much lower in their reading comprehension skills and did not catch up to their typically developing peers. We discuss the implications for children with learning disabilities and the development of their reading comprehension skills.

Correspondence concerning this article should be addressed to Jamie M. Quinn, 2010 Levy Ave., Suite 100, Tallahassee, FL, 32310. Contact: jmquinn@fsu.edu.
Author Notes
Jamie M. Quinn, Florida Center for Reading Research, Florida State University
Richard K. Wagner, Florida Center for Reading Research, Florida State University
Yaacov Petscher, Florida Center for Reading Research, Florida State University
Greg Roberts, The Meadows Center for Preventing Educational Risk, The University of Texas at Austin
Andrew J. Menzel, RiskRecon, LLC
Christopher Schatschneider, Florida Center for Reading Research, Florida State University.

Keywords

Vocabulary knowledge; reading comprehension; learning disabilities; latent change score modeling

Reading comprehension is a vital skill for successful educational attainment, yet only 36% of 4th grade students and 34% of 8th grade students in the United States are proficient in reading comprehension according to a national reading assessment (National Assessment of Educational Progress, 2015). Five to 17 percent of the population of children are identified as having a learning disability (LD), approximately 80% of which have specific problems with word reading and comprehension of written text (Shaywitz & Shaywitz, 2003). Students who exhibit severe difficulties with reading in the primary grades are likely to continue to struggle with reading throughout school (Foorman, Francis, Shaywitz, Shaywitz, & Fletcher, 1997) and are at a higher risk for dropping out of school, experiencing unemployment, and being incarcerated (Newman et al., 2010).

Reading comprehension is a complex process, but is largely driven by two necessary but not singularly sufficient elements: word reading and linguistic comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990). For children with an LD, deficits in either component can lead to problems with reading comprehension. Difficulties with phonological awareness can lead to deficits in decoding, a hallmark of dyslexia or reading disabilities (Vellutino & Fletcher, 2005), and these decoding deficiencies affect a student's ability to comprehend written text (Gough & Tunmer, 1986). Students who have normal decoding skills, but have poor language skills or specific language impairments, are also more likely to have problems with reading comprehension (Catts, Adlof, & Ellis-Weismer, 2006; Catts, Fey, Tomblin, & Zhang, 2002).

Although decoding is a stronger predictor of reading comprehension in the early years of elementary school when reading instruction begins (e.g., Hoover & Gough, 1990; Ouellette & Beers, 2010; Tilstra, McMaster, Van den Broek, Kendeou, & Rapp, 2009; Verhoeven & van Leeuwe, 2008), language skills, particularly vocabulary knowledge, become increasingly important to reading comprehension in the late elementary years in both monolingual and bilingual populations (Geva & Farnia, 2012; Kim & Wagner, 2015; Kim, Wagner, & Lopez, 2012; Ouellette & Beers, 2010; Van Gelderen, Schooen, Stoel, De Glopper, & Hulstijn, 2007; Verhoeven & van Leeuwe, 2008, 2012).

Relations between Vocabulary Knowledge and Reading Comprehension

According to the report of the National Reading Panel (NRP), "reading comprehension… cannot be understood without examining the critical role of vocabulary learning and instruction and its development." (p. 4–1, National Institute of Child Health and Human Development, 2000). Research has supported vocabulary knowledge as one of the best predictors of reading comprehension (e.g., Beck & McKeown, 1981; Cain & Oakhill, 2011; McKeown, Beck, Omanson, & Perfetti, 1983; Quinn, Wagner, Petscher, & Lopez, 2015; Torgesen, Wagner, Rashotte, Burgess, & Hecht, 1997), such that, "vocabulary provides the foundation for grammatical knowledge, definitional vocabulary, and listening comprehension." (NELP, 2008, p. 75). There are a few plausible hypotheses for the basis of this relation between vocabulary and reading comprehension.

Aptitude hypothesis

Anderson and Freebody (1979) posited that vocabulary knowledge and reading comprehension share a joint association with verbal aptitude. According to this hypothesis, children who perform better on vocabulary tests do so because their verbal aptitude is better, and verbal aptitude is implicated in better text comprehension (Anderson & Freebody, 1981). To this end, the best indicator of verbal ability is typically vocabulary knowledge, and verbal ability in turn predicts growth in reading comprehension (Oakhill, Cain, & Bryant, 2003; Sattler, 2001; Stage, Abbott, Jenkins, & Berninger, 2003; Sternberg & Powell, 1983; Woodcock, McGrew, & Mather, 2007).

Instrumental hypothesis

A second explanation is that one must understand the words in a passage to grasp the intended meaning of that passage (Anderson & Freebody, 1979, 1981). Anderson and Freebody named this the *instrumental hypothesis*—that is, vocabulary knowledge is instrumental to successful comprehension of written text. Indeed, children with a smaller vocabularies are more likely to have difficulties with text comprehension and recall (Perfetti, Landi, & Oakhill, 2005). Relatedly, Perfetti's (2007) *lexical quality hypothesis* posits that when the lexical representation of a word is high in quality, the reader can more quickly access semantic information about that word, expediting their comprehension of written words in context through more fluent and automatic reading (Perfetti & Hart, 2001).

Knowledge hypothesis

Anderson and Freebody (1979) suggested that in addition to an instrumental and an aptitude relation, there is a third hypothesis, the knowledge hypothesis, whereby vocabulary and reading comprehension are said to be indirectly related through general knowledge or background knowledge on the topic. Background knowledge is directly implicated in reading comprehension, as having prior knowledge on the topic of the text is useful for drawing inferences and imputing missing information (Anderson & Pearson, 1984). Nagy (2007) extended this knowledge hypothesis and proposed that vocabulary and reading comprehension are correlated by virtue of their joint association with metalinguistic awareness. Morphological awareness is a specific form of metalinguistic awareness that facilitates acquisition of new vocabulary words (Carlisle, 2007; Spencer et al., 2015; Tighe & Schatschneider, 2015). Broadly, metalinguistic awareness facilitates comprehension monitoring-the active process of monitoring one's comprehension of written text- and comprehension monitoring has a direct effect on reading comprehension (Cain, Oakhill, & Bryant, 2004; Perfetti, Marron, & Foltz, 1996; Wagoner, 1983).

Reciprocal hypothesis

The previously mentioned theories posit that vocabulary influences reading comprehension, but the reciprocal hypothesis states that individual differences in reading comprehension may in turn influence vocabulary growth (Elleman, Lindo, Morphy, & Compton, 2009).

Being better at reading comprehension may help not only in inferring the meanings of new words but also in sharpening and extending the meanings of already known words. Indeed, much of a child's vocabulary is acquired by inferring the meanings of words encountered in conversation and in context as opposed to being directly taught and learned (e.g., Cain, 2007; Cunningham & Stanovich, 1991; Nagy, Herman, & Anderson, 1985; Nagy & Scott, 2000; Perfetti et al., 2005; Sternberg, 1987). Through multiple exposures of words in spoken and written forms, children incrementally gain deep knowledge of the definitions and usages of those words (Elleman et al., 2009). Children are thus better situated to read complex passages, which in turn increases text exposure, which further deepens and broadens a child's lexicon of known words through a phenomenon known as the "Matthew Effect" (Stanovich, 1986).

Longitudinal Relations between Vocabulary Knowledge and Reading Comprehension

After controlling for the effects of word identification skills, intelligence, linguistic skills, and inference-making skills, the relation between vocabulary knowledge and reading comprehension holds across the early elementary grades (e.g., Muter, Hulme, Snowling, & Stevenson, 2004; Verhoeven, van Leeuwe, & Vermeer, 2011). Results from path-analytic studies that included autoregressive effects (Gollob & Reichardt, 1987) indicated that vocabulary knowledge predicted performance on reading comprehension tasks independently of previous levels of reading comprehension (de Jong & van der Leij, 2002; Verhoeven & van Leeuwe, 2008; Verhoeven et al., 2011). Correlations between vocabulary and reading comprehension from Kindergarten to grade 2 are moderate when measured with receptive tasks ($.38 \leq rs \leq .44$) and moderate to large when measured with expressive tasks ($.53 \leq rs \leq .70$; c.f. Ouellette, 2006). In addition to a longitudinal instrumental relation, research has supported reciprocal longitudinal relations between vocabulary knowledge and reading comprehension after accounting for word decoding between first and third grade (Verhoeven et al., 2011) and after accounting for self-regulation between first to second grade (Connor et al., 2016). From grade 1 to grade 2, the correlation between vocabulary knowledge and reading comprehension increased from $r = .34$ in fall of grade 1 to $r = .51$ spring of grade 2 (Connor et al.), consistent with findings from other longitudinal studies of elementary school children (e.g., Ouellette, 2006; Quinn, Wagner, Petscher, Lopez, 2015; Reynolds & Turek, 2012).

The previously reported studies used cross-lagged panel modeling (Verhoeven et al., 2011)—a method which has been criticized (Hamaker, Kuiper, & Grasman, 2015)—and structural equation modeling (Connor et al., 2016)—which does not account for individual differences in growth (Ferrer & McArdle, 2010; McArdle, 2009). Latent change score (LCS) modeling combines the advantages of cross-lagged panel models and parallel process growth models to explain both within person and between person differences in the co-development of related processes (Ferrer & McArdle, 2010; McArdle, 2009).

LCS modeling was recently used to support the unidirectional instrumental hypothesis (Quinn et al., 2015; Reynolds & Turek, 2012), such that vocabulary was a leading indicator of change in reading comprehension. LCS modeling was also used to estimate bidirectional relations between verbal IQ and reading (e.g., Ferrer et al., 2007) and between full-scale IQ and reading (Ferrer et al. 2010), supporting a modified reciprocal aptitude hypothesis. In the latter study, Ferrer et al. (2010) discovered differential reciprocal relations regarding students with and without reading problems. The authors suggested that these differential results may have occurred because struggling readers tend to read less than their typically developing peers, and are therefore exposed to fewer vocabulary words and world knowledge through repeated reading. The authors also posited that readers with dyslexia are at a large disadvantage at such an early grade because there is a dissociation between their reading and their IQ, or that a reciprocal relation could have existed between factors outside of the scope of their study.

The Development of Reading in Students with Learning Disabilities

In 2013, 5.7 million children received special education services under Individuals with Disabilities Education Act (IDEA; DePaoli et al., 2015). With an average of 10% of the population having a learning disability, it is imperative we understand any differences in the development and growth of language and literacy skills for children with and without an LD.

Developmental differences between typical students and students with an LD

Students with an LD have more difficulties than typically developing students on measures of reading (Wanzek, Otaiba, & Petscher, 2014). In their large scale study of over a hundred thousand students in second and third grade, Wanzek et al. (2014) estimated that of students identified as having a learning disability ($n = 10{,}339$), over 50% of these students scored below the 20th percentile on a measure of reading comprehension, compared to just 17% of the typically developing students. Additionally, students with an LD started with lower oral reading fluency skills and made less progress over time, never catching up to their typically developing peers (Wanzek et al.).

There are two developmental theories that may underlie why children with an LD show differential growth in reading from their peers without an LD: the developmental lag model and the developmental deficit model. The developmental lag model posits that children with an LD start out lower in reading than their peers without an LD, but grow much faster in their reading to catch up over time (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Stanovich, Nathan, & Zolman, 1988). In early longitudinal work, Satz and colleagues (1978) hypothesized that developmental dyslexia is a result of lagging brain maturation, thus slowing the process of learning to read by way of differential delays in the development in the skills necessary for sufficient reading. Reading growth can also be accelerated for students with a language-based difficulties when language interventions are provided early (e.g., Skibbe et al., 2008), so it would follow that instruction focused on language components, specifically grammar, syntax, and vocabulary, would have a positive impact on reading comprehension for struggling students (NELP, 2009).

The developmental deficit model posits that children with an LD, especially an LD that is reading-based, are poor readers because of a deficit in a skill that does not properly develop (Cromer, 1970; Francis et al., 1996). Candidate skills might be related to their word reading skills (e.g., phonological awareness or decoding; Gough & Tunmer, 1986) or related to their linguistic comprehension skills (e.g., vocabulary knowledge or background knowledge, Gersten, Fuchs, Williams, & Baker, 2001). According to recent work, Spencer and colleagues (2018; 2019) discovered that for children with specific reading comprehension difficulties, or those students with reading comprehension problems although they have appropriate decoding skills, their oral language weaknesses were best indicated as a developmental lag rather than a developmental deficit. This is an important distinction, as a developmental deficit might be an important factor in the development of reading problems (Spencer & Wagner, 2018; Spencer, Wagner, & Petscher, 2019).

The Present Study

Investigating the co-development of language and literacy is especially important for a special population such as students with a school-identified LD. In a study on the environmental effects of poverty for children 1 or 2 years old, by age 3, children from wealthier families knew more words than children from middle class families, and knew even more words than children from low-income families (Hart & Risley, 1995). If children from low income families start preschool knowing fewer words and these children are more likely to have a learning disability, and if we know that vocabulary knowledge is vital to the development of reading comprehension, it is imperative to study the co-development of these skills in a sample of students, both with and without an LD, from a high-risk population of proportionally lower income students.

With the reauthorization of IDEA in 2004, the Department of Education set out to fund low-performing schools with at-risk students through the Reading First initiative. This initiative was designed to improve the reading performance of students who were at risk for reading problems based largely on higher levels of poverty. In concordance with Reading First, states introduced progress monitoring networks to track the progress and development of important early academic skills necessary for educational attainment. Further, states also updated the identification procedures for classifying students with learning disabilities.

The present study sought to examine the co-development of vocabulary knowledge and reading comprehension in a sample from largely Reading First schools using a progress monitoring database from Florida in the US. These at-risk schools have a larger sample of students school-identified as having a learning disability, as rates of LD identification are higher in households with lower socioeconomic status (Shifrer, Muller, & Callahan, 2011). We used multiple group latent change score modeling to investigate the co-development of vocabulary knowledge and reading comprehension from kindergarten through fourth grade for students with and without a school-identified LD. The LCS models helped to answer the following research questions:

1. Is vocabulary knowledge a leading indicator of change in reading comprehension (in support of a instrumental hypothesis)?
2. Is reading comprehension a leading indicator of change in vocabulary knowledge (in support of a reverse instrumental hypothesis)?

3. Are there bidirectional relations between vocabulary knowledge and reading comprehension (in support of a bidirectional instrumental hypothesis)?
4. Or, do vocabulary knowledge and reading comprehension grow together, and there are no direct influences (in support of an aptitude or knowledge hypothesis)?
5. Are there differential developmental relations for students with an LD relative to their typically developing peers? Can these differential relations be described as a developmental lag or developmental deficit?

Method

Participants

The analyses in this study were performed on secondary data. The participants for this study (N = 14,773, 51% female) came from 440 primarily lower-performing schools with Title I designations from 47 districts in Florida, United States. Data were collected from the 2003–2004 school year (Kindergarten) to the 2007–2008 school year (fourth grade) and were recorded in a large state database housed at the Florida Center for Reading Research called the Progress Monitoring and Reporting Network (PMRN). The PMRN database was created to monitor the performance of students in Reading First schools in concordance with the 2004 reauthorization of the Individuals with Disabilities Education Act (IDEA). Reading First was a federally funded initiative designed to improve the reading performance of students who were at risk for reading problems based largely on higher levels of poverty. Although the database included primarily Reading First schools, a small number of non-Reading First schools opted to be included in the database (however, we do not have access to information on which of the 440 schools were or were not designated as Reading First). The total sample demographics were diverse: 51% female, 49.0% White, 26.5% Black, 18.7% Hispanic, 1.4% Mixed ethnicity, 4.2% Asian/Pacific Islander, and 0.3% American Indian. Sixty-five percent of the students in year 1 of data collection were eligible for free or reduced-price lunch or attended a USDA-approved Provision 2 school, which provided meals for free to eligible students.

LD sample—Students were school-identified as having an LD in concordance with the reauthorization of the Individuals with Disabilities Education Act in 2004. Adopted in September 2004, Rule 6A-6.0331 of the Florida Administrative Codes stated that for students in kindergarten through grade twelve, the following activities must be completed prior to referral for learning disabilities: (1) conferences between the caretakers and teacher to address problem areas, (2) at least two observations of student behavior by separate people, (3) evaluation of medical and educational records, (4) review of attendance records, (5) screening for vision, speech, language, and hearing to rule out sensory deficits, and (6) two attempts at intervention (i.e. increased instruction, change in schedule, etc.). If a student required further evaluation after meeting the above criteria, the referral process began by writing a formal request for evaluation of a student's eligibility for education services. After parental consent was given, the evaluation began with standardized tests validated for the intended use of diagnosis, and no single assessment was used as the sole criterion for eligibility for extended services.

The sample from the present study were from a new cohort of students that started Kindergarten during the 2003–2004 school year. Because these students belong to over 400 distinct schools from Florida, it was impossible to determine how quickly or how consistently the new IDEA 2004 criteria were applied when identifying students as having a learning disability. In order to have a better chance at proper state-wide implementation of the new statutes, we used the 2006–2007 LD identification status as the student's classification in the present analysis.

Using the school-based indicator from the database, four percent of the students ($n = 627$) were school identified as having a learning disability in Grade 4. The ethnicities of these students were similar to the total sample: 48.2% White, 27.0% Black, 20.3% Hispanic, 1.0% Mixed ethnicity, 3.0% Asian/Pacific Islander, 0.6% American Indian; however the percentage of male students was larger (67.1% male in the LD sample versus 48% male in the non-LD sample), and a larger proportion of children were eligible for free or reduced-price lunch (77.1% eligible versus 65% for the typically developing children).

Measures

One measure each of vocabulary knowledge and reading comprehension were administered in the spring of each school year (between the middle of February to the end of the school year).

Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997)—This measure tests breadth of receptive vocabulary knowledge, and requires pointing to one of four pictures that represents the target word. Scores were reported as standard scores with a mean of 100 and a standard deviation of 15. Alternative form reliability exceeds 0.88; criterion-related validity coefficients with reading ability range from 0.69 to 0.91 (Williams & Wang, 1997). The PPVT was administered individually. Data were available annually from Kindergarten through 3rd grade.

Stanford Achievement Test – 9: Reading Comprehension (SAT-9; Harcourt Brace Educational Measurement, 1996)—The SAT-9 measured reading comprehension through the two Reading Comprehension subtests from the Primary 1 and Primary 2 forms. These forms administered to first grade students and second grade students, respectively. Primary 1 includes riddles with missing components that a child uses pictures to solve, modified cloze-style short passages, and a reading passage with questions for inferencing. Primary 2 assess reading comprehension through passages that increase in length and difficulty, and include progressively more difficult questions about inferencing and knowledge. Internal consistency reliability exceeds 0.87, concurrent validity coefficients with a state-mandated test of reading comprehension range from 0.78 to 0.84 (Florida Department of Education, 2007). The SAT-9 was administered in groups; data were available for 1st and 2nd grade. The scores on the SAT-9 are presented as scaled scores, a conversion of the students' raw scores, where a difference in scores is the same across the range of abilities and is comparable across ages and grades.

Stanford Achievement Test – 10: Reading Comprehension (SAT-10; Harcourt Brace Educational Measurement, 2003)—In 2005, school districts switched to the 10^{th} version of the SAT to address the out-of-date norms of the previous version. SAT-10 contains the same format of questions for the reading comprehension portion of the SAT-9, but includes more valid and reliable items based on the latest norms of the K-12 population for that period of time. Internal consistency and validity coefficients were similar to the SAT-9 (Florida Department of Education, 2007). The SAT-10 was administered in groups; data were available for 3^{rd} and 4^{th} grade. The scores on the SAT-10 are presented as scaled scores, a conversion of the students' raw scores, where a difference in scores is the same across the range of abilities and is comparable across ages and grades.

Methodology

Bivariate latent change score (LCS) models—Bivariate latent change score (LCS) models make use of longitudinal data to define the nature of dynamic relations between two variables over time. In bivariate LCS models, change is informed by three components (see Online Supplemental Materials I, Figure S1a, for a generic figure). The first component serves as a function of constant yearly change ($\alpha * \eta$), where α is set to 1 for each measurement to assume equally spaced occasions of measurement, and η is the resulting estimated yearly change coefficient. The second component represents the change proportional to the previous level of that variable, representing the autoregressive nature of the models (β). The third component represents change relative to the previous level of the second variable, or the cross-lagged relation (γ). This cross-lagged parameter estimates whether one construct is a leading indicator of change in the second construct. These three parameters may each independently account for variance in the latent change scores, which represent how much change is expected from time t to time $t + 1$. Additionally, each parameter has an associated variance component to determine if there are individual differences in growth and change.

Change scores are calculated as the combination of the above three sources of individual differences in change:

$$\Delta y_{ti} = \left(\alpha_y * \eta_{1yi}\right) + \left(\beta_y * y_{(t-1)i}\right) + \left(\gamma_{yx} x_{(t-1)i}\right), \text{ and}$$
$$\Delta x_{ti} = \left(\alpha_x * \eta_{1xi}\right) + \left(\beta_x * x_{(t-1)i}\right) + \left(\gamma_{xy} y_{(t-1)i}\right)$$

where Δy_{ti} refers to the change in construct y for individual i at time t (for a detailed description, see McArdle, 2009; Petscher, Quinn, & Wagner, 2016).

Multiple model constraints are imposed for both invariance testing and model identification purposes. Measurement error terms are set to equivalence to establish longitudinal measurement invariance. The single-indicator factor loadings for each latent variable at each time point are set to 1 for model identification purposes. The models also assume equally spaced measurement occasions. These models were fit using Mplus 8 (Muthén and Muthén, 1998–2017).

Competing bivariate LCS models were fit to the data for students without an LD and for students with an LD. Four models were tested in each group to determine the developmental dynamics of vocabulary and reading comprehension: 1) A bidirectional cross-lagged model, where vocabulary is posited to be a leading indicator of change in reading comprehension and reading comprehension is posited to be a leading indicator of change in vocabulary, 2) a unidirectional model where reading comprehension was a leading indicator of change in vocabulary knowledge, 3) a second unidirectional model where vocabulary was a leading indicator of change in reading comprehension, and 4) a model where there were no cross-lagged influences between reading comprehension and vocabulary. Models 1, 2, and 3 posited alternative views of the instrumental hypothesis, such that vocabulary is instrumental in improving reading comprehension (Models 1 and 3), and that reading comprehension is instrumental in improving vocabulary knowledge (Models 1 and 2). Model 4 tested the hypothesis that the intercepts and slopes of vocabulary and reading comprehension were correlated, but there were no cross-lagged direct influences between the two variables, analogous to a parallel-process growth model, and most closely related to the aptitude or knowledge hypothesis.

Group differences—Once the dynamics of growth were determined for the groups in separate LCS models, a multiple group analysis was performed in Mplus 8. These models estimated any differences in the growth within and relations between vocabulary knowledge and reading comprehension. Two models were compared with and without constraints on the auto-regressive parameters, cross-lagged parameters, intercepts, slopes, and residual variances. These constraints were tested using MODEL TEST in Mplus, which computes a Wald Test on the constrained parameters and provides a statistic compared against a chi-squared distribution.

Nesting of Students

Students were nested in classrooms, classrooms were nested within schools, and schools were nested within school districts in the state of Florida. Accounting for cross-classification when children move across clustering units is especially complex due to the longitudinal nature of these data. The Huber-White sandwich estimator was chosen as a method to cluster-correct model standard errors in Mplus. This procedure uses an upper-level clustering unit (we chose the student's school at the 2003–2004 measurement of PPVT as the clustering unit) to correct for heteroscedastic errors or disturbances. This nesting procedure was chosen because of the aforementioned difficulty in estimating cross-classified data with longitudinal measurement, and because the prediction of upper-level variances was outside of the scope of the present study.

Model Fit and Model Comparison

Model fit was assessed using the Satorra-Bentler adjusted robust Chi-Square test of model fit (χ^2), the Root Mean Squared Error of Approximation (RMSEA) and its 90% confidence interval, the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), and the sample-sized adjusted Bayesian Information Criteria (nBIC). Lower values of the Chi-Square test with a non-significant *p*-value are preferred, however; given the large sample size used in this study, we expected large values of this fit statistic with significant *p*-values. This metric was considered in combination with the other model fit statistics, such that models with higher values of the CFI and TLI (>.90 preferred) and lower values of the RMSEA (<.08

preferred) and the nBIC (closer to negative infinity) were preferred and were used to establish better fitting models (Kline, 2016).

To compare nested models, the chi-square difference test was computed using the Satorra-Bentler adjusted chi-square and the scaling correction factor. In certain cases, the MLR difference test can result in a negative, uninterpretable value. Where necessary, the MODEL TEST command in Mplus 8 was used to compute a Wald Test statistic to test nested models with negative chi-square difference tests. This Wald Test constrains hypothesized pathways to zero and computes a statistic with associated degree(s) of freedom that are compared against a chi-square distribution.

Data Management

Missing data—Missing data were handled with Mplus 8 (Muthén and Muthén, 19982017) using full-information maximum likelihood (FIML; Anderson, 1957) estimation. This procedure has been found to outperform other *ad hoc* procedures such as list-wise and pair-wise deletion when handling missing data (Enders & Bandalos, 2001).

Staggered time points and z-scoring—Prior to fitting the models, measures were z-scored using a method similar to Ferrer et al. (2007) to place the constructs on the same scale in order to improve the interpretability of parameter estimates. Because vocabulary knowledge was measured starting in Kindergarten, and reading comprehension was measured starting in first grade, z-scores were centered at first grade for both measures and the resulting parameter estimates reflect this approach to standardization (See Table 1). Further, PPVT scores were originally entered as standard scores in the dataset, and no raw scores were available. We calculated estimated raw scores per grade level with a function in R version 3.3.3 (R Core Team, 2017) by reversing the raw to standard score conversion table in the PPVT norms booklet. As multiple raw scores have the same equivalent standard score, particularly at the upper and lower tails (e.g. raw scores between 156 to 204 at age 6 receive a standard score of 160), we approximated those with standard scores that covered a range of raw scores as having the average of those raw scores (e.g., 160 standard score became 180 raw score). We anticipated this method was valid, given that it is difficult to achieve a score considered to be at the ceiling or floor of the PPVT in this age range. The estimated PPVT raw scores were then converted to developmental *z*-scores to place the comprehension and vocabulary measures on the same scale. The resulting model parameters are interpretable as standardized units of measurement.

Results

Sample Statistics

Correlations and descriptive statistics including means, standard deviations, and ranges for the measured variables are presented in Table 1. Tables containing these sample statistics broken down by students with and without an LD are presented in Table 2 and Table 3. The SAT scaled scores indicated that students performed in the average range and experienced positive change in reading comprehension over time. Standard scores for vocabulary knowledge indicate the sample scored similarly to national averages, as these scores were standardized with a mean of 100 and a standard deviation of 15. Accordingly, the developmentally scaled z-scores for PPVT show an increase in vocabulary knowledge over time. The standard deviations and minimum and maximum values suggest considerable individual differences across years and measures. Correlations were significant and positive, ranging from moderate between measures to high within measures (*r*s = .354 – .755).

Comparing Table 2 and Table 3, there are clear differences in the scale scores for SAT 9/10 and the standard scores for PPVT between students with and without a learning disability. Beginning with the scale scores for SAT, students without an LD scored an average of 566 in first grade, as compared to an average of 509 for students with an LD. This score difference of about 40–50 points persisted at all four time points. For the standard scores of PPVT, students without a LD scored at the normative average (96 – 101) at all time points. However, for students with an LD, they scored about 10 points below the normative average (88 – 91) from kindergarten through third grade. The skewness and kurtosis values are largely comparable and are between 0 and 1, indicative of normally distributed variables. The two exceptions are Grade 3 PPVT for the non-LD students (kurtosis = 1.67) and for Kindergarten PPVT for the LD students (kurtosis = 1.45), where these distributions are slightly leptokurtic.

Latent Change Score Modeling

We fit four bivariate LCS models to each sample to determine the functional forms of growth and dynamic relations between vocabulary and reading comprehension: 1) a bidirectional model, where vocabulary was a leading indicator of change in reading comprehension and reading comprehension was a leading indicator of change in vocabulary, 2) a unidirectional model (reading comprehension was a leading indicator of change in vocabulary knowledge), 3) a second unidirectional model (vocabulary as a leading indicator of change in reading comprehension), and 4) a model where there were no cross-lagged influences between reading comprehension and vocabulary.

Non-LD Sample—The bivariate, bidirectional model fit the data well (See Table 4; Model 1A, $\chi^2(22) = 1232.58$, $p < .001$, *RMSEA* = .062 (90% *CI* .059 − .065), *CFI* = .984, *TLI* = .980, *SRMR* = .06, *nBIC* = 176600). Due to the large sample size ($n = 14{,}140$), the chi-square test of model fit was significant and large, but all other fit statistics indicated a good fitting model. Removing the cross-lagged pathways from vocabulary knowledge to change in reading comprehension (by constraining the pathway to zero) significantly degraded model fit (Model 1B as compared to Model 1A: negative Satorra-Bentler $\Delta\chi^2$; *Wald* (1) = 24.53, $p < .001$). In addition, removing the pathways from reading comprehension to change in vocabulary knowledge (by constraining the estimate to zero) resulted in a poorer fitting model (Model 1C as compared to Model 1A: Satorra-Bentler $\Delta\chi^2(1) = 90.23$, $p < .001$, *Wald*(1) = 53.53, $p < .001$). Removing all cross-lagged pathways resulted in the worst fitting model (Model 1D, as compared to Model 1A: Satorra-Bentler $\Delta\chi^2(2) = 54.92$, $p < .001$, *Wald*(2) = 55.63, $p < .001$). The results indicated that there were bidirectional influences between vocabulary knowledge and reading comprehension for typically developing students, and that these leading influences existed after accounting for the effects of the slope parameters and the auto-regressive change parameters.

LD Sample—The bivariate, bidirectional model fit the data excellently (See Table 4; Model 2A, $\chi^2(22) = 49.60$, $p < .001$, *RMSEA* = .045 (90% *CI* .028 − .061), *CFI* = .98, *TLI* = .97, *SRMR* = .04, *nBIC* = 7682.38). However, removing the cross-lagged pathways from vocabulary knowledge to change in reading comprehension (Model 2B) did not result in a significant change in model fit (Model 2B as compared to Model 2A, Satorra-Bentler $\Delta\chi^2(1) = 0.30$, $p = .583$; *Wald*(1) = .713, $p = .398$). Similarly, removing the cross-lagged pathways from reading comprehension to change in vocabulary knowledge did not degrade model fit (Model 2C as compared to Model 2A, Satorra-Bentler $\Delta\chi^2(1) = 1.69$, $p = .194$; *Wald*(1) = .231, $p = .631$). A model where all cross-lagged pathways were removed was not significantly worse fitting than the original bidirectional model (Model 2D as compared to Model 2A, Satorra-Bentler $\Delta\chi^2(2) = 0.68$, $p = .411$; *Wald*(2) = .757, $p = .685$). These pathways were eliminated in the final model for model parsimony. For students with an LD, above and beyond the effect of the constant slope and proportional change parameters, there were no significant direct cross-lagged pathways between vocabulary knowledge and reading comprehension (Model 2D preferred: $\chi^2(24) = 47.27$, $p < .001$, *RMSEA* = .039 (90% *CI* .022 − .056), *CFI* = .99, *TLI* = .98, *SRMR* = .04, *nBIC* = 7677.31). For students with an LD, vocabulary knowledge and reading comprehension grew together, but did not have reciprocal or direct relations over time.

Multiple group model—After fitting individual group models to determine the functional forms of growth and dynamic relations, a multiple group model was fit that imposed measurement invariance and parameter invariance across children with and without an LD. Using the dynamic relations from the separate models (bidirectional influences for students without an LD, no coupling for students with an LD), the first model imposed strict measurement invariance (constrained latent and observed intercepts and residual variances), constrained the means and variances of the intercepts and slopes in vocabulary and reading comprehension across groups, and constrained the covariances between intercepts and slopes across groups. This model fit the data poorly ($\chi^2[64] = 2728.86$, $p < .001$, *RMSEA* = .075 [90% *CI* .073 − .078], *CFI* = .95, *TLI* = .96, *SRMR* = .12, *nBIC* = 185927.10).

In the second model, the following parameters were freed for the LD group: the proportional change parameters for vocabulary and reading comprehension ($df = 2$), the means of the intercepts and slopes ($df = 4$), the variances of the intercepts and slopes ($df = 4$), and the covariances between intercepts and slopes ($df = 6$). This final model fit the data well, $\chi^2(48) = 1341.75$, $p < .001$, *RMSEA* = .060 (90% *CI* .058 − .063), *CFI* = .98, *TLI* = .97, *SRMR* = .06, *nBIC* = 184337.98. Model estimated intercepts, slopes, growth parameter covariances, and auto-regressive pathways were significantly different for students with an LD compared to those students without an LD. The parameters of each model are presented in Table 5 and are discussed further below. Separate path diagrams for each group are presented in the Online Supplemental Materials II (see Figure S2a and Figure S2b).

Model Estimated Change Score Equations—The model-estimated growth parameters are presented in Table 5. However, the individual parameters from the model are only part of an equation for estimating individual differences in change. The parameters must be interpreted within the context of these change score equations. Conveniently, these change score values (within rounding errors) can be extracted from the output file of the final model by requesting the Technical 4 output in Mplus. One can also compute by hand the change score equations using the model-estimated parameters.

The final change score equation for vocabulary knowledge for students without an LD was computed as:

$$\Delta PPVT_{NoLD} = 0.89 + \left(-0.03 * PPVT_{t-1,\,NoLD}\right) + \left(-0.11 * SAT_{t-1,\,NoLD}\right).$$

For students with an LD, the equation was:

$$\Delta PPVT_{LD} = 0.70 + \left(-0.11 * PPVT_{t-1,\,LD}\right) + \left(0 * SAT_{t-1,\,LD}\right).$$

The equations for change in vocabulary knowledge show that for each year, the predicted latent change score increased by 0.89 standard deviation units for children without an LD, and 0.70 standard deviation units for children with an LD. Change proportionally decreased relative to previous PPVT scores for children without an LD (−0.03) and to a larger degree for students with an LD (−0.11). Change in PPVT scores also decreased proportionally by 0.11 points relative to the SAT score at the previous time point, but only for students without an LD. This does not mean that SAT negatively predicts PPVT; its interpretation is that with increasing scores on the SAT, its contribution to growth in PPVT decreases, such that children with higher SAT change less in vocabulary knowledge than students with lower SAT.

The final change score equation for reading comprehension for students without an LD was:

$$\Delta SAT_{NoLD} = 1.05 + \left(-0.60 * SAT_{t-1,\,NoLD}\right) + \left(0.17 * PPVT_{t-1,\,NoLD}\right).$$

For students with an LD was:

$$\Delta SAT_{LD} = 0.65 + \left(-0.38 * SAT_{t-1,\,LD}\right) + \left(0 * PPVT_{t-1,\,LD}\right).$$

The average slope parameter for children without an LD (1.05) was one third of a standard deviation higher than the slope parameter for students with an LD (0.65). Change in reading comprehension proportionally decreased relative to previous reading comprehension scores for children without an LD (−0.60) and with an LD (−0.38). Previous PPVT scores positively predicted change in reading comprehension for students without an LD, but this parameter was not significantly different from zero for students with an LD.

Since the individual parameters in a LCS model should not be interpreted in isolation (as seen above with the negative cross-lagged parameter for SAT to change in PPVT for students without an LD), we can calculate an average change score with a given starting value. Given a child without an LD who has average vocabulary and average reading comprehension scores at Kindergarten, the following change is estimated:

$$\text{No LD}\Delta\text{Voc}_{K \rightarrow 1} 0.92 = 0.89 + (-0.03 * -0.91) + (-0.11 * 0)$$

An individual, who is average at vocabulary in Kindergarten, increases by nearly a full standard deviation in their vocabulary knowledge between Kindergarten and first grade (provided they do not have an LD). The other change equations for the following years are calculated as follows when the average scores for each variable are entered:

$$\text{No LD}\Delta\text{Voc}_{1 \rightarrow 2}: 0.88 = 0.89 + (-0.03 * 0.01) + (-0.11 * 0.05)$$
$$\text{No LD}\Delta\text{Voc}_{2 \rightarrow 3}: 0.76 = 0.89 + (-0.03 * 0.84) + (-0.11 * 0.98)$$

Over time, vocabulary knowledge gains decelerate but remain positive and strong from Kindergarten through third grade. However, for children with an LD, the following equations are estimated:

$$\text{LD}\Delta\text{Voc}_{K \rightarrow 1}: 0.80 = 0.69 + (-0.11 * -1.03) + (0 * 0.00)$$
$$\text{LD}\Delta\text{Voc}_{1 \rightarrow 2}: 0.71 = 0.69 + (-0.11 * -0.20) + (0 * -1.21)$$
$$\text{LD}\Delta\text{Voc}_{2 \rightarrow 3}: 0.64 = 0.69 + (-0.11 * 0.44) + (0 * -0.12)$$

Children with an LD start with slightly lower average scores in vocabulary (−1.03 versus −0.9 for students without an LD), and they tend to have slightly lower change scores than an average child without an LD. Students without an LD make more gains on average in the vocabulary breadth from Kindergarten to third grade.

For reading comprehension, the following change scores were estimated for children without an LD who had average vocabulary and reading comprehension scores:

$$\text{No LD}\Delta\text{RC}_{1 \rightarrow 2}: 1.02 = 1.05 + (-0.63 * 0.054) + (0.20 * 0.01)$$
$$\text{No LD}\Delta\text{RC}_{2 \rightarrow 3}: 0.60 = 1.05 + (-0.63 * 0.978) + (0.20 * 0.84)$$
$$\text{No LD}\Delta\text{RC}_{3 \rightarrow 4}: 0.32 = 1.05 + (-0.63 * 1.678) + (0.20 * 1.62)$$

Reading comprehension gains decelerated quickly over time from a one standard deviation increase between first grade and second grade to about one third of a standard deviation increase between third grade and fourth grade. For children with an LD who have average vocabulary and reading comprehension scores:

$$\text{LD}\Delta\text{RC}_{1 \rightarrow 2}: 1.11 = 0.65 + (-0.38 * -1.21) + (0 * -0.20)$$
$$\text{LD}\Delta\text{RC}_{2 \rightarrow 3}: 0.70 = 0.65 + (-0.38 * -0.12) + (0 * 0.44)$$
$$\text{LD}\Delta\text{RC}_{3 \rightarrow 4}: 0.38 = 0.65 + (-0.38 * 0.72) + (0 * 1.15)$$

Children with an LD had higher estimated change scores between all three sets of grades. However, these students start over one standard deviation below the mean of reading comprehension of students without an LD. They do not “catch up” to their peers by the time the reach Grade 4.

Growth trajectories and split violin plots—The results of the models indicated that students without an LD started significantly higher in reading comprehension (about 1 SD higher) and somewhat higher (about 0.1 SD higher) in vocabulary knowledge. According to the estimated change scores, students without an LD grew somewhat more in vocabulary but grew somewhat less in reading comprehension than students with an LD. To visualize these differences, Figure 1 presents split violin plots generated with ggplot2 (Wickham, 2009). These plots present average growth lines as estimated by the LCS model and show split distributions of scores for a randomly selected subset of 200 students (100 per group). For vocabulary knowledge (Figure 1, top), students with or without an LD started with similar levels of vocabulary. The average growth lines split somewhat over time, but the distributions of vocabulary scores largely overlap at each time point.

For reading comprehension, there were noticeable differences in the intercepts and slopes between groups (Figure 1, bottom). Students with an LD (light blue) started over a standard deviation lower in reading comprehension on average in grade 1. Further, the average growth lines were far apart, and did not converge by fourth grade. There was little overlap between the distributions of scores for the students without an LD (salmon colored side) and students with an LD (light blue colored side).

The estimated growth trajectories for those same randomly selected students (100 students without an LD and 100 students with an LD) were created in ggplot2 (Wickham, 2009) and are presented in Figure 2. The salmon colored lines, representing students without an LD, are more towards the top of the figure, indicative of higher scores in reading comprehension. More light blue lines are represented in the bottom half of the figure, indicative of lower scores in reading comprehension. Over time and on average, the children represented in this figure who have an LD do not catch up to their peers in reading comprehension. Students with and without an LD have similarly shaped growth trajectories, indicative of relatively similar gains in vocabulary knowledge over time.

In order to describe average 1-year change in both vocabulary and reading comprehension depending on starting values, presented in Figure 3 is a vector plot grouped by LD status: expected growth for students without an LD are represented in salmon, expected growth for students with an LD are represented in light blue. Also pictured are actual estimates of vocabulary and reading comprehension skills for the randomly selected subset of students. According to the overlap across the x-axis, students with and without an LD largely overlap with their vocabulary knowledge. However, students without an LD extend farther up the y-axis of the figure, indicative of higher average reading comprehension scores. The light blue and salmon colored ellipses represent the 90% confidence ellipse for each group, such that 90% of students within each group should fall into those ranges. The salmon colored ellipse ranges farther up the y-axis than the light blue ellipse, i.e., students with an LD are more likely to have problems with reading comprehension. Regarding the differences in trajectory lines, for children who start with high vocabulary knowledge but low reading comprehension skills, students who do not have an LD are expected to grow nearly twice as much in reading comprehension as compared to students who have an LD (see the bottom right portion of Figure 3). For most points in the figure, students without an LD have more change than students without an LD. For the points where students with an LD are said to outgain students without an LD (e.g., the top left portion of the graph), it is unlikely that a student with an LD will score in this range (see 90% confidence ellipse).

Discussion

The present study investigated the developmental relations between vocabulary knowledge and reading comprehension in children with and without a school-identified learning disability using multiple group bivariate LCS modeling. The first goal was to investigate the developmental dynamics between vocabulary knowledge and reading comprehension separately for students with and without an LD. The second goal was to determine whether the developmental dynamics and growth of reading comprehension and vocabulary are different in students with and without a learning disability. We discuss results relevant to each goal in turn.

Developmental Influences between Vocabulary Knowledge and Reading Comprehension

The results of the LCS models supported bidirectional influences between vocabulary knowledge and reading comprehension over time, consistent with Ferrer et al. (2007). However, this effect was only present for children without a learning disability. For the students without an LD, we attribute our detecting the significant effect of reading comprehension on subsequent development of vocabulary when other studies have not (Ferrer et al., 2010; Quinn et al., 2015; Reynolds & Turek, 2012) to the greater sensitivity afforded by our large sample size, which reduced the standard errors of model estimates and increased our power to detect smaller effects.

From a methodological perspective, the bidirectional influences found with LCS modeling would not have been detected using latent growth curve modeling, including a parallel process growth curve model that modeled growth in both constructs simultaneously. LCS models have a parallel process growth curve model embedded in them, which allows for testing of bidirectional influences between two constructs above and beyond the influence of correlated intercepts and slopes and proportional change parameters.

Does the development within and between reading comprehension and vocabulary knowledge differ between students with and without a learning disability?

Yes. There were large differences in intercept values of reading comprehension between children with an LD and typically developing children. Typically developing children on average began over a standard deviation higher on reading comprehension $\left(\mu_{R_{NoLD}} = 0.04\right)$ than their peers with an LD$\left(\mu_{R_{LD}} = -1.24\right)$. The intercept values for vocabulary were similar for students without an LD $\left(\mu_{V_{NoLD}} = -0.91\right)$ and students with an LD $\left(\mu_{V_{LD}} = -1.07\right)$. The slopes for reading comprehension and for vocabulary were higher for students without an LD $\left(\eta_{R_{NoLD}} = 1.05, \eta_{V_{NoLD}} = 0.89\right)$ than for students with an LD $\left(\eta_{R_{LD}} = 0.65, \eta_{V_{LD}} = 0.69\right)$.

There were bidirectional cross-construct influences for students without an LD, such that vocabulary was a leading indicator of change in reading comprehension, and reading comprehension was a leading indicator of change in vocabulary. For children with an LD, there were no significant cross-lagged pathways, indicating that although these constructs have correlated growth, there are no direct, instrumental relations between vocabulary and reading comprehension. One reason for the lack of a bidirectional influence is that students with an LD typically bring less of their vocabulary knowledge to text comprehension tasks, and their reading comprehension growth lags behind (Gersten, Fuchs, Williams, & Baker, 2001).

Developmental Lag versus Developmental Deficit?

Our results indicated that students not only started lower in reading comprehension when they had a school-identified LD, but they also did not catch up to their typically developing peers in reading comprehension by grade 4. This is indicative of a development deficit, as they were not nearing the average reading comprehension skills of their typically developing peers. However, reading comprehension change scores for students with an LD were somewhat higher than for students without an LD, indicating that our model may have looked at too narrow of a developmental period for a catch-up effect to manifest.

Lower performing students with a school-identified LD—The change score equations from the results section used average levels of reading comprehension and vocabulary for each group of students. When the equations for change in reading comprehension are calculated again with reading comprehension scores one standard deviation below the mean (using the mean minus the standard deviation from Table 3), it results in the following change scores:

$$\begin{aligned} 1.41 &= 0.65 + (-0.38 * -2.00) + (0 * -0.20) \\ 1.01 &= 0.65 + (-0.38 * -0.94) + (0 * 0.44) \\ 0.70 &= 0.65 + (-0.38 * -0.13) + (0 * 1.15) \end{aligned}$$

Children with an LD who start with *worse* reading comprehension grow *more* than students with an LD who start with *average* reading comprehension. Although the population of students with an LD on average do not catch up to the population of students without an LD on average, lower performing students with a school-identified LD may eventually catch up to their higher-performing peers with a school-identified LD. This gain may be in vain, however, if students with an LD still lag behind their typically developing peers throughout primary and secondary school. It is also unclear if these lowest performing students with a school-identified LD had the largest gains in reading comprehension because of regression to the mean or if these students had more targeted instruction or were in specialized intervention programs. Because we do not have detailed information on the educational contexts of these students, we were unable to answer this question.

The Importance of Educational Contexts and Home Environment

We studied development as it occurred in the context of the instructional practices that were used in the schools from which our participants were drawn. We have no indication of curriculum, classroom, or district level differences in instructional practices or learning environments. Instructional practices can particularly influence performance, especially in the early elementary years (Cameron, Connor, & Morrison, 2005; Creemers & Reezigt, 1966; Rimm-Kaufman, La Paro, Downer, & Pianta, 2005). The results of a developmental study such as ours do not imply what would happen if either construct were targeted for more intense instruction or intervention. We also cannot determine how classroom, school or district-level differences affected the learning of students with school-identified learning disabilities.

Further, our results indicated that students with an LD started much lower than their peers without an LD on our test of reading comprehension. These students did not catch up to their typically developing peers by fourth grade. This same pattern of reading struggles has been shown in students described as "garden variety" low achievers (Francis et al., 1996; Gough & Tunmer, 1986; Share, 1996; Stanovich, 1988). This similarity between students with an LD and garden-variety poor readers begs the question, "Do students with an LD and garden variety poor readers differ in ways that matter within a learning or educational context?" We are not able to determine whether there are discernible differences between these subpopulations of students from the present study, but a previous meta-analysis indicated that students with an LD have more pronounced difficulties with reading than do struggling readers not identified as LD (Fuchs, Fuchs, Mathes, & Lipsey, 2000). Knowing from the results of our study that a student with an LD starts with very low levels of reading comprehension on average, these models may help to inform the literature on garden-variety struggling readers at the lowest end of the distribution as compared to their peers without reading struggles.

In addition to classroom contexts, it is important to also consider the impact of home environments on language and literacy outcomes. Hart and Risley's seminal investigation (1995) on the stark class differences in vocabulary exposure found that children in higher-income households learned far more words due to more overall exposure and through rich conversations with their caretakers. Within the present study, the sample chosen from Title-I schools had a large proportion of children eligible for free- and reduced-price lunch. Although this metric is not an accurate portrayal of students who come from low income families, it serves as a useful proxy measure in this case. Of students eligible for FRL in Kindergarten, these students scored 16 standard points *lower* on the PPVT (mean = 87) compared to students who were not eligible for FRL (mean = 103), regardless of LD status. Table 2 and 3 also show this difference for students with an LD, whereby they scored 12 points below students without an LD on average in Kindergarten and remain behind through third grade. A further breakdown of FRL status versus LD status reveals striking trends (see Table S1 in supplemental materials): students with and without an LD have much lower scores entering Kindergarten when they are eligible for free or reduced-price lunch in their school relative to similar peers who are ineligible for FRL.

Limitations and Future Directions

Third variable relations—For students without an LD, we found similar parameter estimates of the leading influence of vocabulary knowledge on reading comprehension as compared to previous studies (e.g., Reynolds & Turek, 2012) and found the additional leading influence of reading comprehension to changes in vocabulary knowledge as seen in Ferrer et al. (2007). However, other important predictors of these relations were not tested, which should be considered a limitation of our study. For example, an effect of vocabulary knowledge on reading comprehension might be mediated by word decoding skills. Decoding is both a critical determinant of reading comprehension (see García and Cain, 2014, for a meta-analysis on the relations between decoding and reading comprehension), and may be related to vocabulary knowledge during this developmental period (e.g., Chiappe, Chiappe, & Gottardo, 2004; see also the *lexical restructuring model*, Metsala & Walley, 1998). It is easier for a child to decode real words than it is non-words, especially if these words are in the child's spoken lexicon (Lonigan, 2007; Lonigan, Burgess, & Anthony, 2000; Verhoeven, Van Leeuwe, & Vermeer, 2011), and word decoding has been found to predict the development of vocabulary knowledge in later elementary years (Verhoeven et al. 2011).

In addition to decoding, reading fluency is increasingly important to reading comprehension over the elementary years, especially fluency for reading connected text (Geva & Farnia, 2012; Kim, 2015a; Little, Hart, Quinn, Tucker-Drob, Taylor, & Schatschneider, 2017; Petscher et al., 2016). Further, vocabulary knowledge is highly correlated with and can be influenced by other oral language skills such as grammatical knowledge in both monolingual populations (e.g., Kim, 2015b; Uccelli, Galloway, Barr, Meneses, & Dobb, 2015) and in bilingual populations (Hoff, Quinn, & Giguere, 2018). Our results do not rule out variables that could have mediated, moderated, or accounted for the present cross-construct relations (for students without an LD) or absent cross-construct relations (for students with an LD).

Limitations of measurement—Our measure of vocabulary knowledge is a widely used standardized measure of receptive vocabulary skills with a large validation range of ages. Even so, that this measure was the only one available can be seen as a limitation. The National Early Literacy Panel (2008) reported that measures of receptive vocabulary are only weakly correlated with reading comprehension outcomes ($r = .25$), and that vocabulary knowledge is more likely to be an important indicator of vital oral language skills such as grammatical and syntactical knowledge (NELP, 2008). The results of our study reported that PPVT in Kindergarten was correlated with SAT-9 in Grade 1 at $r = .36$. This correlation was smaller for children with a school-identified LD ($r = .28$), but is comparable to the correlation described in the NELP report. Over time, the concurrent correlation between PPVT and SAT increased for children without a school-identified LD (grade 1 $r = .41$; grade 2 $r = .52$; grade 3 $r = .55$) and for children with a school-identified LD (grade 1 $r = .23$; grade 2 $r = .30$; grade 3 $r = .48$).

Regarding our chosen measure of reading comprehension, the SAT-9/10 included increasingly difficult questions starting with the Primary 1 form in Grade 1. This form used three formats that featured increasingly longer texts. It began with two-sentence riddles in which a child selects a picture that best describes the riddle, then shifted to short modified cloze passages with one to three words missing, and finally finished with a short reading passage with multiple choice questions. The criticism with cloze-style comprehension tests is that they are more dependent on word decoding and are less related to oral language, especially in the younger years (Keenan, Betjemann, & Olson, 2008). The increasing correlations between PPVT and SAT-9/10, which evolves in its task style through Primary 1 and Primary 2 and beyond, is evident for the decreasing role of decoding and the increasing role of language in the comprehension of text passages.

Methods for identifying students with learning disabilities—Identification methods for classifying students with learning disabilities have been heavily criticized. There are excellent, recent publications with in-depth discussions of the existing problems with these classification methods, with response to intervention (RTI) methods appearing to be the most reliable methods available (Fletcher & Miciak, 2019; Maki & Adams, in press). With the reauthorization of IDEA in 2004, a discrepancy between IQ and ability can no longer be the sole method for identification purposes, a method which has also been heavily criticisms as inaccurate and problematic (Francis et al., 2005; Stuebing et al., 2002).

As stated in the methods section, the present study included a new cohort of students starting Kindergarten during the 2003–2004 school year, right as IDEA 2004 was being passed by the US Department of Education and being passed and implemented through state-level governments. It is not possible to see if the new criteria were accurately and consistently being applied in the identification of students with an LD. In using Grade 4 identification status, we had the best chance to have enough time for state-wide implementation of the new guidelines. However, it may be possible these new standards had not yet been properly implemented across all of the school districts, so the results need to be considered limited in regards to proper identification of students with an LD. Still, these students were far behind their peers without a school-identified LD with regards to reading comprehension initial status and growth.

Our decision to compare students with and without a school-identified LD instead of using researcher-based criteria typically seen in the literature might be viewed as an additional limitation. One reason we decided to focus on school-identified students was that these were the students who were theoretically receiving specialized services in the form of Individualized Education Programs (IEPs) or other services aimed at ensuring their educational success.

Another reason we used school-defined LD status was that researchers have suggested that when identifying students with an unobservable, latent trait such as having a learning disability, it is imperative to use multiple measures to avoid inaccurate and invalid identification due to measurement error (Fletcher & Miciak, 2019; Francis et al., 2005). Although school-identification procedures have been criticized, these methods are largely based on multiple criteria. Researcher-based identification using single measures can have poor overlap, leading to inaccurate identification. Keenan and colleagues (2014) showed that only half of the time does a student who performs poorly on a measure of reading comprehension (i.e., below the 25th percentile) also perform poorly on another measure of reading comprehension. This poor overlap has been shown in multiple publications, such as the poor overlap seen between students identified using researcher-based criteria and students with school-identified LDs (Quinn & Wagner, 2015), low longitudinal stability of classifying students with a reading disability (Schatschneider, Wagner, Hart, & Tighe, 2016), and poor agreement among differing definitions of response to intervention (Barth et al., 2008).

As a test of our own data, we looked at the level of overlap between students with a school-identified LD and students who were below the 20th percentile in SAT-9/10 reading comprehension (see Online Supplemental Materials V for details and additional cross-tabulations). Among the 627 students with a school-identified LD, 364 of those students were below the 20th percentile on SAT 9 in Grade 1 and 366 were below the 20th percentile on SAT-10 in Grade 4. That corresponds to about one-half of the students who met researcher-based criteria and who were also school-identified as having a learning disability. The children identified using researcher-based criteria were largely different from the students identified using school-based criteria (*kappa* estimates of agreement ranged from $.041 \leq \kappa \leq .283$). Students below the 20th percentile on SAT who did not have a school-identified LD maintained lower reading scores than SLD-only students and had lower vocabulary scores in Kindergarten and maintained lower average vocabulary scores through 3rd grade (See Online Supplemental Materials VI). As a model comparison, a multiple group LCS model was fit comparing the sample of students with an LD (n = 627) to a sample of students persistently below the 20th percentile on SAT who did not have an LD (n = 331; see Online Supplemental Materials VII). The persistently poor readers started with slightly lower initial mean status in RC, but started a half standard deviation below the mean of vocabulary ($\mu_{0v} = -1.57$) relative to students with an LD ($\mu_{0v} = -1.07$). The model also showed that both students with an LD and the persistently low achievers did not have significant cross-lagged co-development, such that there were no direct relations between vocabulary knowledge and change in their reading comprehension or vice versa. The students with an LD had a similar co-developmental pattern for their vocabulary knowledge and reading comprehension as compared to students who could be considered “garden variety poor readers” (Stanovich, 1988); however, these persistently poor readers also started with lower vocabulary knowledge than students with an LD. Further research is needed to understand the differences and similarities in the co-development of these important literacy skills in poor readers and students with learning disabilities.

Sample limitations—Additional limitations of the study are that the sample, although large and relatively representative of the Southeastern United States, was drawn primarily from Title 1 schools with a larger portion of at-risk youth and our sampling was limited to the early to middle elementary grades. We also did not have the opportunity to disaggregate data according to Title 1 schools compared to non-Title I schools. For the students in the present sample during this developmental period (K- 4th grade), it was clear that vocabulary knowledge is important for the development of reading comprehension skills. However, the results could be different for other populations of students, especially students with more advanced in reading skills. Previous research has suggested that learning vocabulary words or inferring meaning from context is an easier task for more skilled comprehenders than it is for less skilled comprehenders (e.g., Cain, Oakhill, & Lemmon, 2004).

Relatedly, although 80% of students with learning disabilities have specific struggles with reading (Shaywitz & Shaywitz, 2003), we could not disaggregate our sample to determine if the students school-identified as having an LD had problems with reading or if they had general struggles with learning not specific to reading. These students on average scored very low on reading comprehension compared to their typically developing peers, but there was a large range of ability in year 1 (scale score range: 437 – 643). Future studies should consider outcomes for children with differing kinds of learning disabilities.

The sample of school-identified students was small (*n* = 627) relative to the sample of typically developing students (*n* = 14,146). As a result, the power to detect significant effects is significantly higher in the non-LD sample of students. The standard errors are much smaller, and the *t*-values of estimates are more likely to be significant as a result. It is unlikely that power explains why there were no bidirectional relations for students with a school-identified LD. For the pathway of reading comprehension to change in vocabulary knowledge, the estimated parameter was quite small (0.036; see Figure S2a in the Supplemental Materials), and was not significantly different from zero (*se* = 0.043, *p* = .399). For the estimated pathway from vocabulary to change in reading comprehension, the pathway parameter is more substantial (−0.301), but the standard error is very large (*se* = 0.631, *p* = .633). Further, the sign of this pathway would suggest that students with larger vocabularies grow *less* in reading comprehension over time

Lastly, our models were estimated using data that was collected through Grade 4. Although we found proximal evidence for a model that supports a developmental deficit in reading comprehension for these students with an LD, we cannot rule out the existence of a developmental lag when our data do not go beyond fourth grade. Because children with lower reading comprehension grow more than children with higher reading comprehension grow, these lower-performing students may "catch up" to their typically developing peers if given enough time. This would be indicative of a developmental lag. Future studies should investigate if these developmental differences persist into middle and high school to determine whether students with an LD will catch up to their typically developing peers.

Implications for students with a school-identified LD in low performing schools

According to our model, the worse a child's starting level of reading comprehension, the more opportunity they have for growth. However, it is unlikely that having larger change scores would help the lowest performing children with an LD in the short run (i.e., through grade 4), as they started two standard deviations lower in reading comprehension in grade 1 than typically developing children with average reading comprehension in grade 1.

With regards to evidence for a developmental delay, it would be important to determine if this developmental delay occurs in the presence of a problem with decoding or in the absence of a problem with decoding. This distinction is important, especially with regards to children with specific reading comprehension problems, as children with comprehension problems despite adequate decoding skills are usually the product of a developmental delay (Spencer & Wagner, 2018; Spencer et al., 2019). For the present sample, we did find support for a developmental delay for both language and comprehension. Since we do not know the decoding skills of the included sample, we were unable to determine if children with an LD had problems with comprehension that were specific to or independent of their decoding. This has implications for instruction and targeted interventions, as students with LD might benefit more from decoding-focused programs, since we did not find evidence that their vocabulary knowledge was related to growth in their reading comprehension skills.

Students with an LD in this sample may have had learning disabilities not specific to reading (as discussed in the limitations section above). It is difficult to disentangle differential growth patterns between different types of learning disabilities, because this disaggregation did not exist in our database. Outside of this fact, our model estimated individual differences in change for both students with and without an LD. Future studies that incorporate interventions for students with learning disabilities should consider methods that assess individual change for progress monitoring purposes.

For students who do not catch up to their peers by the fourth grade, interventions aimed at improving the reading comprehension skills for late elementary and middle school students with learning disabilities or broad struggles with reading have been successful (see Scammacca; Solis, Ciullo, Vaughn, Pyle, Hassaram, & Leroux, 2012; Wanzek, Wexler, Vaughn, & Ciullo, 2010). These interventions were the most impactful for proximal measures of reading comprehension created by the research team, but were less impactful for distal measures such as standardized reading comprehension tests.

There is limited evidence that vocabulary interventions improve vocabulary knowledge and also distally improve reading comprehension. If direct teaching of vocabulary is largely ineffective, or only effective with proximal measures, then most word learning must occur incidentally through exposure to new words in differing contexts (Cunningham & Stanovich, 1991; Nagy, Herman, & Anderson, 1985; Nagy & Scott, 2000). However, it is apparent that when prior exposure to unknown words is high, incidental word learning is increased, and reading comprehension is improved (Jenkins, Stein, & Wysocki, 1984). For students with specific reading comprehension impairments, interventions with specific reading comprehension components and specific oral language components are effective, as well as combined text/oral language programs (Clarke, Snowling, Truelove & Hulme, 2010).

Conclusions

In summary, vocabulary knowledge and reading comprehension exhibited bidirectional influences over time for children without an LD, but there were no cross-lagged influences for children with an LD. Children with an LD started with similar levels of vocabulary knowledge in Kindergarten as compared to their typically developing peers, but started much lower in reading comprehension in Grade 1 and did not catch up to their typically developing peers by Grade 4. The results of the present study demonstrate the importance of studying the co-development of related constructs through using a sophisticated statistical method, and that through using LCS modeling, we found evidence for a developmental delay in reading comprehension, whereby students with an LD do not catch up to their peers by fourth grade.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported in part by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (P50 HD052120; P50 HD052117–07) The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health and Human Development, the National Institutes of Health, or the U.S. Department of Education.

References

Anderson TW (1957). Maximum likelihood estimates for multivariate normal distributions when some observations are missing. Journal of American Statistical Association, 52, 200–203. 10.2307/2280845

Anderson RC, & Freebody P (1979). Vocabulary knowledge (Tech. Rep. No.136). Urbana-Champaign: University of Illinois, Center for the Study of Reading.

Anderson R, & Freebody P (1981). Vocabulary knowledge In Guthrie JT (ED.), Comprehension and teaching: Research reviews (pp. 77–117). Newark, DE: International Reading Association.

Anderson RC, & Pearson PD (1984). A schema-theoretic view of basic processes in reading In Pearson PD, Barr R, Kamil ML, & Mosenthal P (Eds.), Handbook of Reading Research (pp. 255–291). White Plains, NY: Longman.

Barth AE, Stuebing KK, Anthony JL, Denton CA, Mathes PG, Fletcher JM, & Francis DJ (2008). Agreement among response to intervention criteria for identifying responder status. Learning and individual Differences, 18(3), 296–307. 10.1016/j.lindif.2008.04.004 [PubMed: 19081758]

Beck I, Perfetti C, & McKeown M (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. Journal of Educational Psychology, 74, 506–521. 10.1037/0022-0663.74.4.506

Cain K (2007). Deriving word meanings from context: Does explanation facilitate contextual analysis? Journal of Research in Reading, 30, 347–359. 10.1111/j.1467-9817.2007.00336.x

Cain K, & Oakhill J (2011) Matthew effects in young readers: Reading comprehension and reading experience aid vocabulary development. Journal of Learning Disabilities, 44, 431–443. 10.1177/0022219411410042 [PubMed: 21772058]

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(1), 31 10.1037/0022-0663.96.1.31

Cain K, Oakhill J, & Lemmon K (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. Journal of Educational Psychology, 96(4), 671 10.1037/0022-0663.96.4.671

Cameron CE, Connor CM, & Morrison FJ (2005). Effects of variation in teacher organization on classroom functioning. Journal of School Psychology, 43, 61–85. 10.1016/j.jsp.2004.12.002

Carlisle JF (2007). Fostering morphological processing, vocabulary development, and reading comprehension In Wagner R, Muse A & Tannenbaum K (Eds.), Vocabulary acquisition: Implications for reading comprehension (pp. 78–103). New York: Guilford Press.

Catts H, Adlof S, & Ellis-Weismer S (2006). Language deficits in poor comprehenders: A case for the simple view of reading. Journal of Speech, Language, and Hearing Research, 49, 278–293. 10.1044/1092-4388(2006/023)

Catts HW, Fey ME, Tomblin JB, & Zhang X (2002). A longitudinal investigation of reading outcomes in children with language impairments. Journal of speech, Language, and hearing Research, 45(6), 1142–1157. 10.1044/1092-4388(2002/093)

Chiappe P, Chiappe DL, & Gottardo. A. (2004). Vocabulary, context and speech perception among good and poor readers. Educational Psychology, 24, 825–843. 10.1080/0144341042000271755

Clarke PJ, Snowling MJ, Truelove E, & Hulme C (2010). Ameliorating children’s reading-comprehension difficulties: A randomized controlled trial. Psychological Science, 21(8), 1106–1116. 10.1177/0956797610375449 [PubMed: 20585051]

Connor CM, Day SL, Phillips B, Sparapani N, Ingebrand SW, McLean L, … Kaschak MP (2016). Reciprocal Effects of Self-Regulation, Semantic Knowledge, and Reading Comprehension in Early Elementary School. Child Development, 87(6), 1813–1824. 10.1111/cdev.12570 [PubMed: 27264645]

Creemers BPM, & Reezigt GJ (1996). School level conditions affecting the effectiveness of instruction. School Effectiveness and School Improvement, 7, 197–228. 10.1080/0924345960070301

Cromer W (1970). The difference model: A new explanation for some reading difficulties. Journal of Educational Psychology, 61, 471–488. 10.1037/h0030288 [PubMed: 5483823]

Cunningham A, & Stanovich K (1991). Tracking the unique effects of print exposure in children—association with vocabulary, general knowledge, and spelling. Journal of Educational Psychology, 83, 264–274. 10.1037/0022-0663.83.2.264

de Jong P, & van der Leij A (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. Scientific Studies of Reading, 6, 51–77. 10.1207/S1532799XSSR0601_03

Elleman AM, Lindo EJ, Morphy P, & Compton DL (2009). The impact of vocabulary instruction on passage level comprehension of school age children: A meta-analysis. Journal of Research on Educational Effectiveness, 2, 1–44. 10.1080/19345740802539200

Ferrer E, McArdle JJ, Shaywitz BA, Holahan JM, Marchione K, & Shaywitz SE (2007). Longitudinal models of developmental dynamics between reading and cognition from childhood to adolescence. Developmental Psychology, 43, 1460–1473. 10.1037/0012-1649.43.6.1460 [PubMed: 18020824]

Ferrer E, & McArdle JJ (2010). Longitudinal modeling of developmental changes in psychological research. Current Directions in Psychological Science, 19, 149–154. 10.1177/0963721410370300

Ferrer E, Shaywitz BA, Holahan JM, Marchione K, & Shaywitz SE (2010). Uncoupling of reading and IQ over time: Empirical evidence for a definition of dyslexia. Psychological Science, 21, 93–101. 10.1177/0956797609354084 [PubMed: 20424029]

Fletcher JM, & Miciak J (2019). The identification of specific learning disabilities: A summary of research on best practices. Austin, TX: Texas Center for Learning Disabilities Retrieved from: https://www.texasldcenter.org/files/resources/SLDManual_Final.pdf

Florida Department of Education. (2007). Assessment & Accountability Briefing Book: FCAT—School Accountability—Teacher Certification Tests—2007. Tallahassee, Florida: Author Retrieved from: http://www.fldoe.org/core/fileparse.php/7478/urlt/briefingbook07web.pdf

Francis DJ, Fletcher JM, Stuebing KK, Lyon GR, Shaywitz BA, Shaywitz SE (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. Journal of Learning Disabilities, 38, 98–108. 10.1177/00222194050380020101 [PubMed: 15813593]

Francis DJ, Shaywitz SE, Stuebing KK, Shaywitz BA, & Fletcher JM (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. Journal of Educational Psychology, 88, 3–17. 10.1037/0022-0663.88.1.3

García JR & Cain K (2014). Decoding 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. 10.3102/0034654313499616

Gersten R, Fuchs LS, Williams JP, & Baker S (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. Review of Educational Research, 71, 279–320. 10.3102/00346543071002279

Geva E, & Farnia F (2012). Developmental changes in the nature of language proficiency and reading fluency paint a more complex view of reading comprehension in ELL and EL1. Reading and Writing, 25, 1819–1845. 10.1007/s11145-011-9333-8

Grimm KJ, An Y, McArdle JJ, Zonderman AB, & Resnick SM (2012): Recent changes leading to subsequent changes: Extensions of multivariate latent difference score models. Structural Equation Modeling: A Multidisciplinary Journal, 19, 268–292. 10.1080/10705511.2012.659627 [PubMed: 23637519]

Hamaker EL, Kuiper RM, & Grasman RPPP (2015). A critique of the cross-lagged panel model. Psychological Methods, 20, 102–116. 10.1037/a0038889 [PubMed: 25822208]

Harcourt Brace Educational Measurement (1996). Stanford Achievement Test Series (9th ed.). San Antonio, TX: Author.

Harcourt Brace Educational Measurement (2003). Stanford Achievement Test Series (10th ed.). San Antonio, TX: Author.

Hart B, & Risley TR (1995). Meaningful differences in the everyday experience of young American children. Baltimore, MD: Paul H. Brookes Publishing.

Hoff E, Quinn JM, & Giguere D (2018). What explains the correlation between growth in vocabulary and grammar? New evidence from latent change score analyses of simultaneous bilingual development. Developmental Science, 21, e12536 10.1111/desc.12536

Jenkins JR, Stein ML, & Wysocki K (1984). Learning vocabulary through reading. American Educational Research Journal, 21, 767–787. 10.3102/00028312021004767

Keenan JM, Betjemann RS, & Olson RK (2008). Reading Comprehension Tests Vary in the Skills They Assess: Differential Dependence on Decoding and Oral Comprehension. Scientific Studies of Reading, 12(3), 281–300. 10.1080/10888430802132279

Keenan JM, Hua AN, Meenan CE, Pennington BF, Willcutt E, & Olson RK (2014). Issues in identifying poor comprehenders. LAnnee psychologique, 114(4), 753–777.

Kim YS-G (2015a). Developmental, component-based model of reading fluency: An investigation of predictors of word-reading fluency, text-reading fluency, and reading comprehension. Reading Research Quarterly, 50, 459–481. 10.1002/rrq.107 [PubMed: 26435550]

Kim YS (2015b). Language and cognitive predictors of text comprehension: Evidence from multivariate analysis. Child Development, 86(1), 128–144. 10.1111/cdev.12293 [PubMed: 25174258]

Kim YSG, & Wagner RK (2015). Text (oral) reading fluency as a construct in reading development: An investigation of its mediating role for children from grades 1 to 4. Scientific Studies of Reading, 19(3), 224–242. 10.1080/10888438.2015.1007375 [PubMed: 25848201]

Kim YS, Wagner RK, & Lopez D (2012). Developmental relations between reading fluency and reading comprehension: A longitudinal study from Grade 1 to Grade 2. Journal of Experimental Child Psychology, 113, 93–111. 10.1016/j.jecp.2012.03.002 [PubMed: 22726256]

Little C, Hart SA, Quinn JM, Tucker-Drob EM, Taylor J, & Schatschneider C (2017). Exploring the co-development of reading fluency and reading comprehension: A twin study. Child Development, 88, 934–945. 10.1111/cdev.12670 [PubMed: 27859016]

Little RJ (1988). A test of missing completely at random for multivariate data with missing values. Journal of the American Statistical Association, 83, 1198–1202. 10.1080/01621459.1988.10478722

Lonigan CJ (2007). Vocabulary development and the development of phonological awareness skills in preschool children In Wagner R, Muse A & Tannenbaum K (Eds.), Vocabulary acquisition: Implications for reading comprehension (pp. 15–31). New York: Guilford Press.

Lonigan CJ, Burgess SR, & Anthony JL (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent variable longitudinal study. Developmental Psychology, 36, 596–613. 10.1037/00121649.36.5.596 [PubMed: 10976600]

Maki KE, & Adams SR (in press). Specific learning disabilities identification: Do the identification methods and data matter? Learning Disability Quarterly. 10.1177/0731948719826296

McArdle JJ (2009). Latent variable modeling of longitudinal data. Annual Review of Psychology, 60, 577–605. 10.1146/annurev.psych.60.110707.163612

McKeown MG, Beck IL, Omanson RC, & Perfetti CA (1983). The effects of long-term vocabulary instruction on reading comprehension: A replication. Journal of Reading Behavior, 15, 3–18. 10.1080/10862968309547474

Metsala JL, & Walley AC (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability In Metsala JL & Ehri LC (Eds.) Word Recognition in Beginning Literacy (pp. 89–120). Mahwah, NJ: Erlbaum.

Muthén LK, & Muthén BO (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Authors.

Nagy WE (2007). Metalinguistic awareness and the vocabulary-comprehension connection In Wagner RK, Muse A & Tannenbaum K (Eds.), Vocabulary acquisition: Implications for reading comprehension (pp. 53–77). New York: Guilford Press.

Nagy WE, & Herman PA (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction In McKeown MG & Curtis ME (Eds.), The Nature of Vocabulary Acquisition (pp. 19–35). Hillsdale, NJ: Lawrence Erlbaum Associates.

Nagy WE, Herman PA, & Anderson RC (1985). Learning words from context. Reading Research Quarterly, 233–253. 10.2307/747758

Nagy WE, & Scott J (2000). Vocabulary processing In Kamil M, Mosenthal P, Pearson PD, & Barr R (Eds.), Handbook of Reading Research (Vol. 3, pp. 269–284). Mahwah, NJ: Erlbaum.

National Early Literacy Panel (2008). Developing early literacy: Report of the National Early Literacy Panel. Washington, DC: National Institute for Literacy Retrieved from https://lincs.ed.gov/earlychildhood/NELP/NELPreport.html

National Institute of Child Health and Human Development. (2000). Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00–4769). Washington, DC: U.S. Government Printing Office Retrieved from https://www.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf

Newman L, Wagner M, Cameto R, Knokey A-M, & Shaver D (2010). Comparisons across Time of the Outcomes of Youth with Disabilities up to 4 Years after High School A Report of Findings from the National Longitudinal Transition Study (NLTS) and the National Longitudinal Transition Study-2 (NLTS2) (NCSER 2010–3008). Menlo Park, CA: SRI International.

Oakhill J, Cain K, & Bryant PE (2003). The dissociation of word reading and text comprehension: Evidence from component skills. Language and Cognitive Processes, 18, 443–468. 10.1080/01690960344000008

Oakhill J, Cain K, & McCarthy D (2015). Inference processing in children: The contributions of depth and breadth of vocabulary knowledge In O'Brien EJ, Cook AE, & Lorch RF Jr. (Eds.), Inferences during Reading (pp 140–159). Cambridge, UK: Cambridge University Press.

Ouellette G (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. Journal of Educational Psychology, 98, 554–566. 10.1037/0022-0663.98.3.554

Ouellette G, & Beers A (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. Reading and writing, 23(2), 189–208. 10.1007/s11145-008-9159-1

Perfetti CA, & Hart L (2001). The lexical bases of comprehension skill In Gorfien D (Ed.), On the consequences of meaning selection (pp. 67–86). Washington, DC: American Psychological Association.

Perfetti CA, Landi N, & Oakhill J (2005). The acquisition of reading comprehension skill In Snowling MJ & Hulme C (Eds.), The Science of Reading: A Handbook (pp. 227–247). Malden: Blackwell Publishing.

Perfetti CA, Marron MA, & Foltz PW (1996). Sources of comprehension failure: Theoretical perspectives and case studies In Cornoldi C & Oakhill JV (Eds.), Reading comprehension difficulties: Processes and remediation (pp. 137–165). Mahwah, NJ: Erlbaum.

Petscher Y, Quinn JM, & Wagner RK (2016). Modeling the co-development of correlated processes with longitudinal and cross-construct effects. Developmental Psychology, 52, 1690–1704. 10.1037/dev0000172 [PubMed: 27732037]

Qian D (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. Canadian Modern Language Review, 56, 282–308. 10.3138/cmlr.56.2.282

Quinn JM, Wagner RK, Petscher Y, & Lopez D (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. Child Development, 86, 159–175. 10.1111/cdev.12292 [PubMed: 25201552]

R Core Team (2018). R: A language and environment for statistical computing Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Reynolds MR, & Turek JJ (2012). A dynamic developmental link between verbal comprehension-knowledge (Gc) and reading comprehension: Verbal comprehension knowledge drives positive change in reading comprehension. Journal of School Psychology, 50, 841–863. 10.1016/j.jsp.2012.07.002 [PubMed: 23245503]

Rimm-Kaufman SE, La Paro KM, Downer JT, & Pianta RC (2005). The contribution of classroom setting and quality of instruction to children's behavior in kindergarten classrooms. The Elementary School Journal, 105, 377–394. 10.1086/429948

Sattler J (2001). Assessment of children: Cognitive applications (4th ed.). La Mesa, CA: Jerome M Sattler Publisher.

Scammacca N, Roberts G, Vaughn S, Edmonds M, Wexler J, Reutebuch CK, Torgesen JK (2007). Intervention for adolescent struggling readers: A meta-analysis with implication for practice.

Portsmouth, NH: RMC Research Corporation, Center on Instruction Retrieved from http://www.centeroninstruction.org/files/Metaanalysis%20Struggling%20Readers1.pdf

Schatschneider C, Wagner RK, Hart SA, & Tighe EL (2016). Using simulations to investigate the longitudinal stability of alternative schemes for classifying and identifying children with reading disabilities. Scientific Studies of Reading, 20(1), 34–48. 10.1080/10888438.2015.1107072 [PubMed: 26834450]

Share DL (1996). Word recognition and spelling processes in specific reading disabled and garden-variety poor readers. Dyslexia, 2, 167–174. 10.1002/(SICI)1099-0909(199611)2:3<167::AID-DYS167>3.0.CO;2-O

Shaywitz SE, & Shaywitz BA (2003). Dyslexia (specific reading disability). Pediatrics in Review, 24, 147–153. Retrieved from http://pedsinreview.aappublications.org/content/24/5/147.full-text.pdf [PubMed: 12728187]

Shifrer D, Muller C, & Callahan R (2011). Disproportionality and Learning Disabilities: Parsing Apart Race, Socioeconomic Status, and Language. Journal of Learning Disabilities, 44(3), 246–257. 10.1177/0022219410374236 [PubMed: 20587753]

Skibbe LE, Grimm KJ, Stanton-Chapman TL, Justice LM, Pence KL, & Bowles RP (2008). Reading Trajectories of Children with Language Difficulties from Preschool through Fifth Grade. Language, Speech, and Hearing Services in Schools, 39(4), 475–486. 10.1044/0161-1461(2008/07-0016)

Solis M, Ciullo S, Vaughn S, Pyle N, Hassaram B, & Leroux A (2012). Reading comprehension interventions for middle school students with learning disabilities: A synthesis of 30 years of research. Journal of Learning Disabilities, 45, 327–340. 10.1177/0022219411402691 [PubMed: 21490168]

Spencer M, Muse A, Wagner RK, Foorman B, Petscher Y, Schatschneider C, Tighe EL & Bishop MD (2015). Examining the underlying dimensions of morphological awareness and vocabulary knowledge. Reading and Writing, 28(7), 959–988. 10.1007/s11145-015-9557-0 [PubMed: 26273128]

Spencer M, & Wagner RK (2018). The comprehension problems of children with poor reading comprehension despite adequate decoding: A meta-analysis. Review of Educational Research, 88(3), 366–400. 10.3102/0034654317749187 [PubMed: 29785063]

Stage S, Abbott R, Jenkins J, & Berninger V (2003). Predicting response to early reading intervention from verbal IQ, reading-related language abilities, attention ratings, and verbal IQ-word reading discrepancy: Failure to validate discrepancy method. Journal of Learning Disabilities, 36, 24–33. 10.1177/00222194030360010401 [PubMed: 15490889]

Stahl S, & Fairbanks M (1986). The effects of vocabulary instruction: A model-based meta-analysis. Review of Educational Research, 56, 72–110. 10.3102/00346543056001072

Stanovich KE (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. Journal of Learning Disabilities, 21, 590–612. 10.1177/002221948802101003 [PubMed: 2465364]

Stanovich KE, Nathan RG, & Vala-Rossi M (1986). Developmental changes in the cognitive correlates of reading ability and the developmental lag hypothesis. Reading Research Quarterly, 21, 267–283. http://www.jstor.org/stable/747709

Stanovich KE, Nathan RG, & Zolman JE (1988). The developmental lag hypothesis in reading: Longitudinal and matched reading-level comparisons. Child Development, 59, 71–86. 10.2307/1130390

Sternberg RJ (1987). Most vocabulary is learned from context In McKeown M, & Curtis ME (Eds.), The nature of vocabulary acquisition (pp. 89–105). Hillsdale, NJ: Erlbaum.

Sternberg RJ, & Powell JS (1983). Comprehending verbal comprehension. American Psychologist, 38, 878–893. 10.1037/0003-066X.38.8.878

Stuebing KK, Fletcher JM, LeDoux JM, Lyon GR, Shaywitz SE, Shaywitz BA (2002). Validity of IQ-discrepancy classifications of reading disabilities: A meta-analysis. American Educational Research Journal, 39, 469–518. 10.3102/00028312039002469

Tighe EL, & Schatschneider C (2015). Exploring the dimensionality of morphological awareness and its relations to vocabulary knowledge in Adult Basic Education students. Reading Research Quarterly, 50, 293–311. 10.1002/rrq.102

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(4), 383–401. 10.1111/j.1467-9817.2009.01401.x

Torgesen JK, Wagner RK, Rashotte CA, Burgess S, & Hecht S (1997). Contributions of phonological awareness and rapid automatic naming ability to the growth of word-reading skills in second- to fifth-grade children. Scientific Studies of Reading, 1, 161–195. 10.1207/s1532799xssr0102_4

Uccelli P, Galloway EP, Barr CD, Meneses A, & Dobbs CL (2015). Beyond vocabulary: Exploring cross-disciplinary academic-language proficiency and its association with reading comprehension. Reading Research Quarterly, 50(3), 337–356. 10.1002/rrq.104

Vellutino FR, & Fletcher JM (2005). Developmental dyslexia In Snowling MJ & Hulme C (Eds.), The Science of Reading: A Handbook (pp. 362–378). Oxford, UK: Blackwell

Verhoeven L, & van Leeuwe J (2008). Prediction of the development of reading comprehension: A longitudinal study. Applied Cognitive Psychology, 22, 407–423. 10.1002/acp.1414

Verhoeven L, & van Leeuwe J (2012). The simple view of second language reading throughout the primary grades. Reading and Writing, 25, 1805–1818. 10.1007/s11145-011-9346-3 [PubMed: 22923881]

Verhoeven L, van Leeuwe J, & Vermeer A (2011). Vocabulary growth and reading development across the elementary school years. Scientific Studies of Reading, 15(1), 8–25. 10.1080/10888438.2011.536125

Wagoner SA (1983). Comprehension monitoring: What it is and what we know about it. Reading Research Quarterly, 23, 328–346. 10.2307/747392

Wagner RK, & Torgesen JK (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. Psychological Bulletin, 101, 192–212. 10.1037/0033-2909.101.2.192

Wagner RK, Torgesen JK, Laughon P, Simmons K, & Rashotte CA (1993). Development of young readers' phonological processing abilities. Journal of Educational Psychology, 85, 1–20. 10.1037/0022-0663.85.1.83

Wanzek J, Otaiba SA, & Petscher Y (2014). Oral Reading Fluency Development for Children with Emotional Disturbance or Learning Disabilities. Exceptional Children, 80(2), 187–204. 10.1177/001440291408000204

Wanzek J, Wexler J, Vaughn S, Ciullo S (2010). Reading interventions for struggling readers in the upper elementary grades. A synthesis of 20 years of research. Reading and Writing, 23, 889–912. 10.1007/s11145-009-9179-5 [PubMed: 21072128]

Wickham H (2009). ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer-Verlag.

Author Manuscript

Author Manuscript

Author Manuscript

Educational Impact and Implications Statement

The present study suggests that for students with a school-identified learning disability (LD), the amount of vocabulary words a student with an LD knows does not have a direct impact on their growth in reading comprehension skills. This may have implications for language-based interventions of reading problems for these children with LD, as we find little evidence that improvements in language will transfer to their reading comprehension delays. This further emphasizes the importance creating individual education plans as a way to stimulate growth in reading comprehension for students with an LD in at-risk environments.

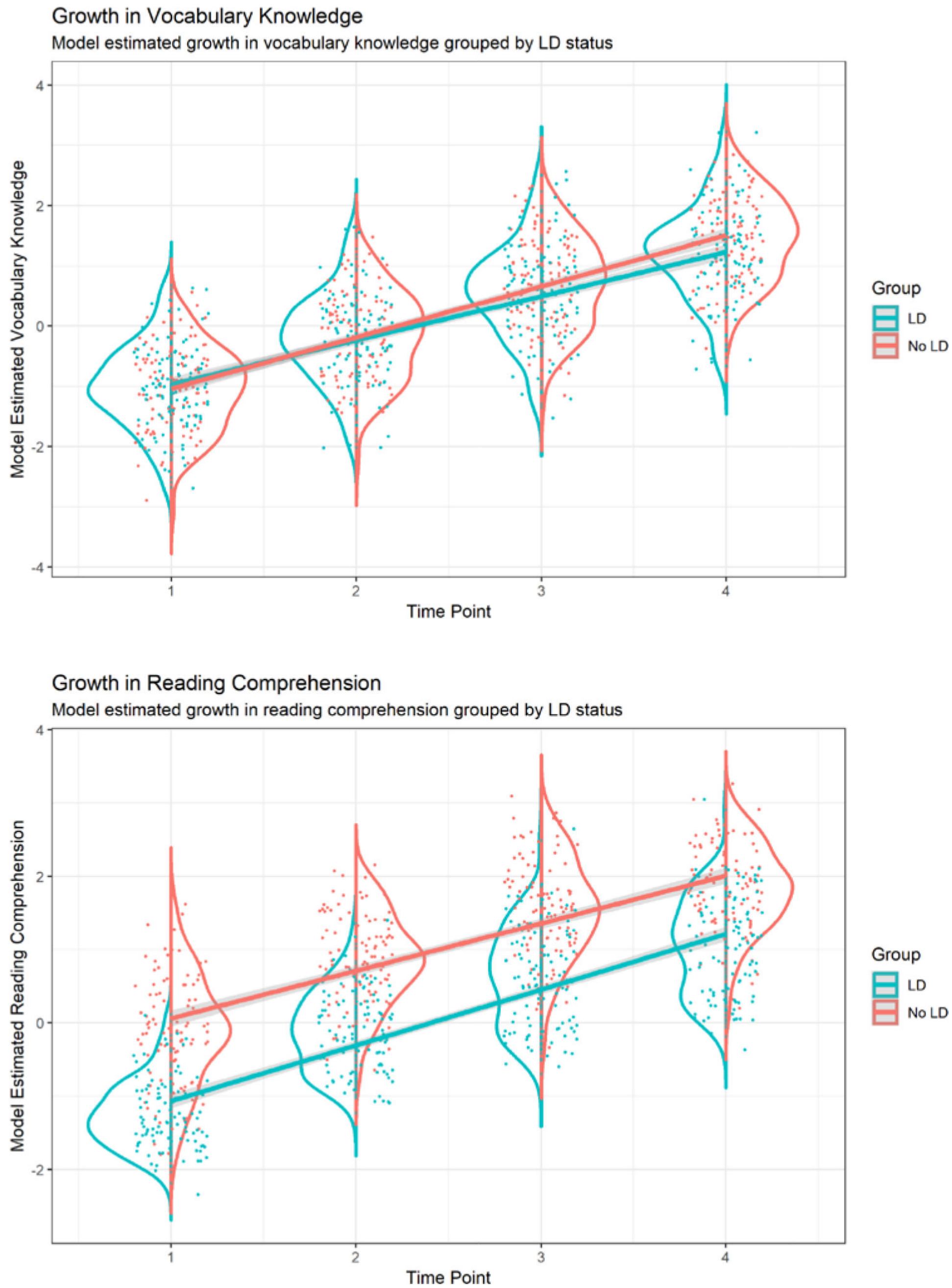

Figure 1.

Split-Violin Plot for Vocabulary Knowledge (top) and Reading comprehension (bottom). This plot shows the distributions for vocabulary knowledge and reading comprehension at each time point for a random subset of 100 students per group. Also pictured is the average growth in each group specified with thick lines. Values on the y-axis are in developmental *z*-scored units scaled at first grade. Light Blue = students with an LD, Salmon = students with no LD.

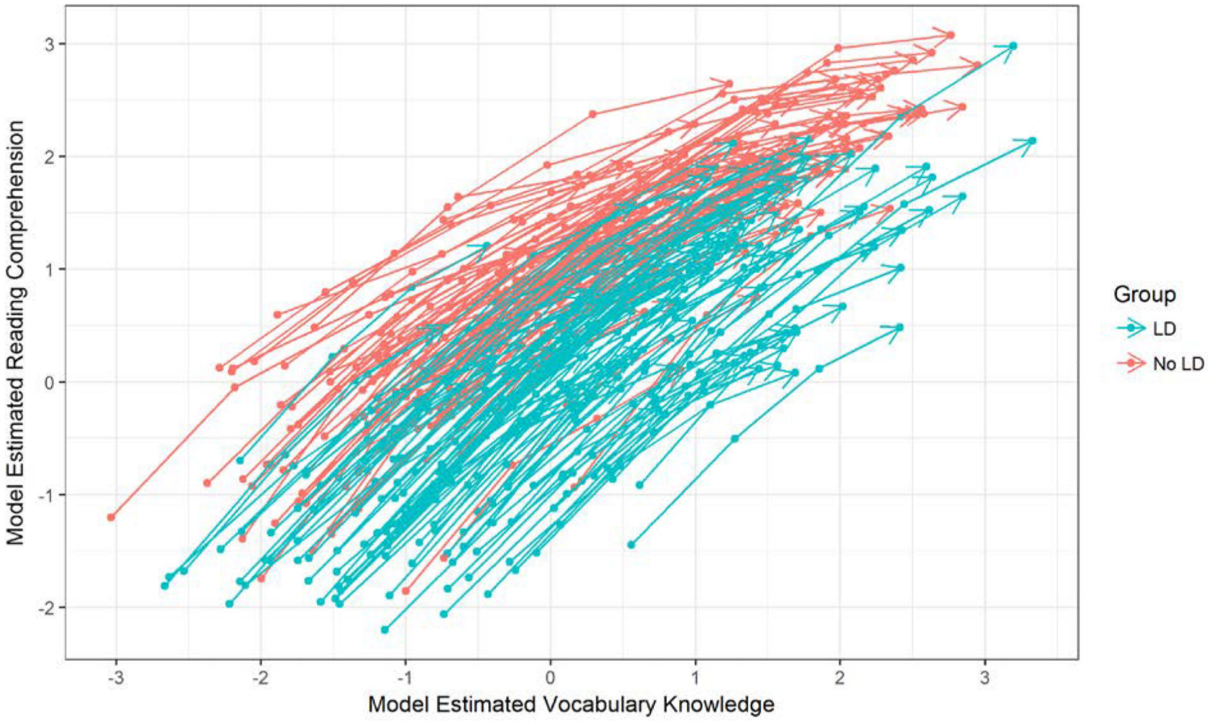

Figure 2.

Estimated developmental trajectories for a random sample of 100 students with LD and 100 students without an LD. This figure shows the expected trajectories estimated by the LCS modeling technique. The LD sample is in light blue, the non-LD sample is in salmon. The four dots on each line represent the four time points.

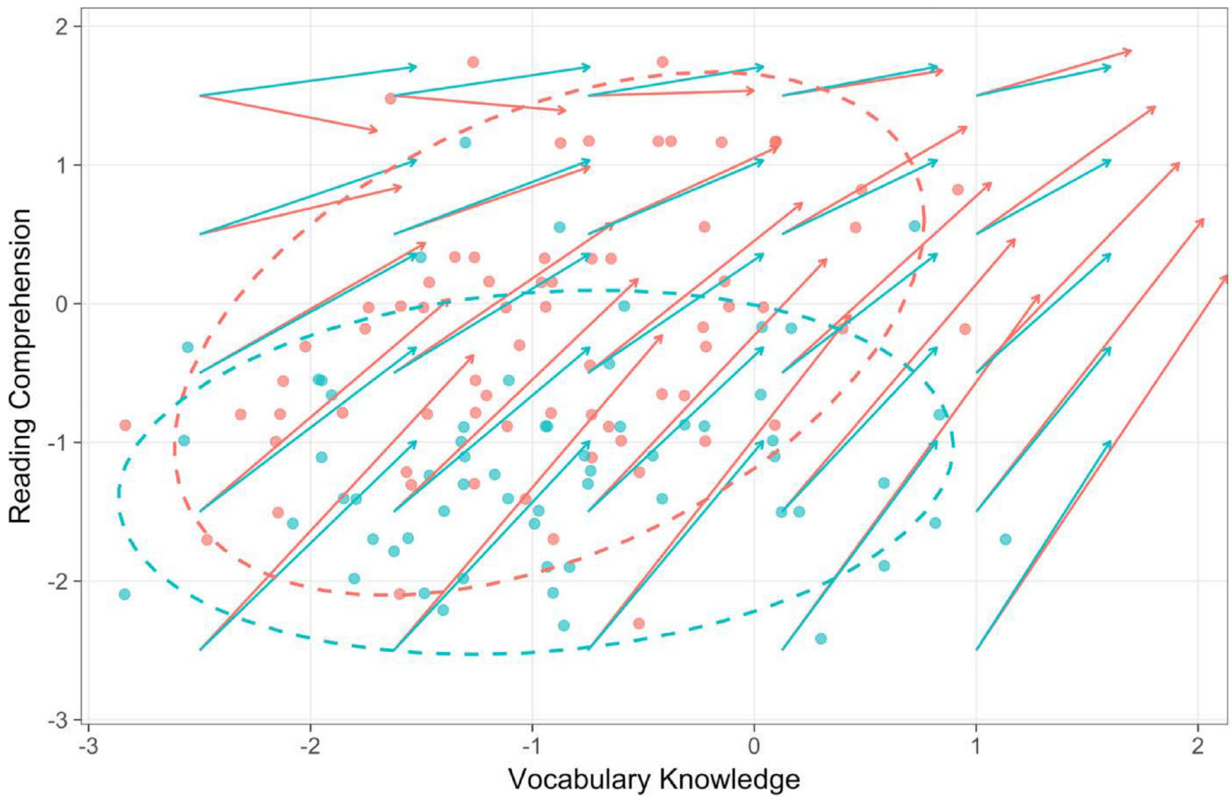

Figure 3.
Vector plot by subgroup. This vector plot shows the expected 1-year growth in reading comprehension and vocabulary knowledge based on a grid of hypothetical starting values, grouped by students with an LD (light blue) and students without an LD (salmon/pink). Overlaid are a random subset of 100 students with an LD (light blue dots) and 100 students without an LD (salmon/pinkish dots). The dotted circles represent the 90% confidence ellipses per group.

Table 1.

Descriptive Statistics and Correlations for the Complete Sample ($n = 14{,}773$)

Measure	1	2	3	4	5	6	7	8
1. Grade 1, SAT-9	1							
2. Grade 2, SAT-9	.65	1						
3. Grade 3, SAT-10	.62	.72	1					
4. Grade 4, SAT-10	.55	.65	.73	1				
5. Grade K, PPVT	.36	.43	.43	.35	1			
6. Grade 1, PPVT	.40	.47	.49	.41	.75	1		
7. Grade 2, PPVT	.44	.51	.54	.46	.70	.76	1	
8. Grade 3, PPVT	.45	.52	.56	.49	.63	.70	.75	1
Scale Scores[a] (SAT-9/SAT-10) or Standard Scores[b] (PPVT-III)								
Mean	564.00	606.45	638.10	651.29	95.88	97.10	98.70	100.59
SD	45.02	36.09	40.52	34.62	13.59	14.17	13.91	13.08
Minimum	415	423	483	497	40	40	40	40
Maximum	667	729	763	776	147	160	160	160
Skewness	0.19	−0.01	0.16	−0.19	−0.09	−0.05	−0.27	−0.43
Kurtosis	−0.30	0.51	0.05	0.35	0.70	0.09	0.44	1.57
Z-scores								
Mean	0	0.97	1.63	1.92	−0.92	0	0.86	1.62
SD	1	0.81	0.90	0.77	0.98	1	1.02	0.97
Minimum	−3.28	−3.11	−1.78	−1.48	−5.10	−4.96	−4.53	−4.41
Maximum	2.27	3.63	4.38	4.67	4.38	5.23	5.60	5.80

Note. SAT-9 = Stanford Achievement Test-9, Reading Comprehension; SAT-10 = Stanford Achievement Test-10, Passage Comprehension. PPVT = Peabody Picture Vocabulary Test. All correlations significant at $p < .01$.

[a]Scale scores were only available for the SAT in the original database.

[b]Standard scores were only available for the PPVT values in the original database.

We used a reverse look-up procedure using the reference tables in the technical manual to estimate raw scores (see Methods section or contact first author for more information on this procedure).

Table 2.

Descriptive Statistics and Correlations for Students without a Learning Disability (*n*=14,146)

Measure	1	2	3	4	5	6	7	8
1. Grade 1, SAT-9	1							
2. Grade 2, SAT-9	.62	1						
3. Grade 3, SAT-10	.60	.70	1					
4. Grade 4, SAT-10	.53	.62	.71	1				
5. Grade K, PPVT	.37	.44	.44	.36	1			
6. Grade 1, PPVT	.41	.48	.50	.42	.75	1		
7. Grade 2, PPVT	.44	.52	.54	.46	.69	.75	1	
8. Grade 3, PPVT	.44	.52	.55	.48	.63	.70	.75	.1
Scale Scores [a] *(SAT-9/SAT-10) or Standard Scores* [b] *(PPVT-III)*								
Mean	566.46	608.42	640.20	653.09	96.20	97.42	99.06	101.00
SD	44.21	34.62	39.29	33.32	13.48	14.04	13.77	12.87
Minimum	415	423	483	497	40	40	40	40
Maximum	667	729	763	776	147	160	160	160
Skewness	0.23	0.11	0.22	−0.10	−0.07	−0.04	−0.27	−0.41
Kurtosis	−0.32	0.49	0.07	0.32	0.66	0.09	0.51	1.67
Z-scores								
Mean	0.05	0.98	1.68	1.96	−0.91	0.01	0.84	1.62
SD	0.97	0.76	0.87	0.73	0.97	1.00	1.01	0.95
Minimum	−3.28	−3.12	−1.78	−1.48	−5.10	−4.96	−4.53	−4.41
Maximum	2.27	3.63	4.38	4.67	2.82	5.23	5.60	5.80

Note. SAT-9 = Stanford Achievement Test-9, Reading Comprehension; SAT-10 = Stanford Achievement Test-10, Passage Comprehension. PPVT = Peabody Picture Vocabulary Test. All correlations significant at $p < .01$.

[a]Scale scores were only available for the SAT in the original database.

[b]Standard scores were only available for the PPVT values in the original database.

We used a reverse look-up procedure using the reference tables in the technical manual to estimate raw scores (see Methods section or contact first author for more information on this procedure).

Table 3.

Descriptive Statistics and Correlations for Students with a Learning Disability (*n*=627)

Measure	1	2	3	4	5	6	7	8
1. Grade 1, SAT-9	1							
2. Grade 2, SAT-9	.54	1						
3. Grade 3, SAT-10	.52	.70	1					
4. Grade 4, SAT-10	.46	.61	.75	1				
5. Grade K, PPVT	.28	.23	.33	.27	1			
6. Grade 1, PPVT	.23	.24	.36	.27	.73	1		
7. Grade 2, PPVT	.23	.30	.42	.32	.74	.76	1	
8. Grade 3, PPVT	.27	.33	.48	.37	.64	.70	.80	1
Scale Scores[a] (SAT-9/SAT-10) or Standard Scores[b] (PPVT-III)								
Mean	509.03	558.59	590.12	610.62	88.26	89.59	89.81	90.66
SD	36.01	37.25	38.46	38.38	14.15	15.00	14.43	14.31
Minimum	437	423	523	517	40	40	58	40
Maximum	643	679	763	728	126	132	128	135
Skewness	0.54	0.36	0.70	0.11	−0.42	−0.13	−0.16	−0.45
Kurtosis	0.19	0.24	0.34	−0.74	1.45	0.14	−0.71	−0.15
Z-scores								
Mean	−1.21	−0.12	0.72	1.03	−1.03	−0.20	0.44	1.15
SD	0.79	0.82	0.85	0.85	0.99	1.05	1.07	1.07
Minimum	−2.80	−3.11	−0.90	−1.04	−5.04	−4.47	−1.89	−2.06
Maximum	1.74	2.53	4.38	3.61	1.50	2.82	2.90	3.76

Note. SAT-9 = Stanford Achievement Test-9, Reading Comprehension; SAT-10 = Stanford Achievement Test-10, Passage Comprehension. PPVT = Peabody Picture Vocabulary Test. All correlations significant at $p < .01$.

[a] Scale scores were only available for the SAT in the original database.

[b] Standard scores were only available for the PPVT values in the original database.

We used a reverse look-up procedure using the reference tables in the technical manual to estimate raw scores (see Methods section or contact first author for more information on this procedure).

Table 4

Model fit statistics for the Bivariate Latent Change Score Models

Model	χ^2	*df*	*SCR*	RMSEA	90% CI	CFI	TLI	SRMR	nBIC
Non-LD Students									
Model 1A	**1232.58**	**22**	**1.35**	**.062**	**.059 –.065**	**0.98**	**0.98**	**0.06**	**176600.34**
Model 1B	1215.92	23	1.42	.061	.058 –.063	0.98	0.98	0.07	176652.80
Model 1C	1312.04	23	1.33	.063	.060 –.066	0.98	0.98	0.06	176674.71
Model 1D	1261.98	24	1.40	.060	.058 –.063	0.98	0.98	0.07	176698.74
LD Students									
Model 2A	49.60	22	1.13	.045	.028 –.061	.98	.97	.04	7682.38
Model 2B	46.56	23	1.22	.040	.023 –.057	.99	.98	.04	7680.12
Model 2C	51.62	23	1.10	.045	.028 –.061	.98	.98	.04	7679.92
Model 2D	**47.27**	**24**	**1.21**	**.039**	**.022 –.056**	**.99**	**.98**	**.04**	**7677.31**

Note. χ^2 = Chi-Square test of model fit; df = degrees of freedom. SCR = scaling correction factor for Satorra-Bentler chi-square difference testing; RMSEA = Root Mean-Squared Error of Approximation. CI = Confidence interval. CFI = Comparative Fit Index. TLI = Tucker-Lewis Index. SRMR = standardized root mean square residual; nBIC = Sample size adjusted Bayesian information criteria.

Table 5.

Parameters for the multiple group LCS model

	Without an LD		With an LD	
Parameter	***Vocab***	***RC***	***Vocab***	***RC***
Initial mean status, $\mu_{intercept}$	-0.91^{***}	0.04^{***}	-1.07^{***}	-1.24^{***}
Linear slope, η_{slope}	0.89^{***}	1.05^{***}	0.69^{***}	0.65^{***}
Auto-Regression, β	-0.03^{**}	-0.60^{***}	-0.11^{**}	-0.38^{***}
Cross-lagged effects, γ				
RC → Δ Vocab	-0.11^{***}		0	
Vocab → Δ RC	0.20^{***}		0	
Variances:				
Intercepts, σ_0^2	0.77^{***}	0.73^{***}	0.79^{***}	0.42^{***}
Slopes, σ_S^2	0.03^{***}	0.10^{***}	0.04^{***}	0.11^{***}
Residual Errors, σ_e^2	0.24^{***}	0.21^{***}	0.24^{***}	0.21^{***}
Covariances:				
Vocab $\mu_{intercept}$ ↔ Vocab μ_{slope}	01^{ns}		$.36^{*}$	
Vocab $\mu_{intercept}$ ↔ RC $\mu_{intercept}$	49^{***}		$.34^{***}$	
Vocab $\mu_{intercept}$ ↔ RC μ_{slope}	$.28^{***}$		$.32^{***}$	
RC $\mu_{intercept}$ ↔ RC μ_{slope}	$.60^{***}$		$.53^{***}$	
RC $\mu_{intercept}$ ↔ Vocab μ_{slope}	$.53^{***}$		$.19^{*}$	
RC $\mu_{intercept}$ ↔ Vocab μ_{slope}	$.51^{***}$		$.51^{***}$	

Note. Vocab= PPVT vocabulary; RC = SAT reading comprehension.

ns = not significant

*** = $p < .001$

** = $p < .01$

* = $p < .05$.

Residual variances were fixed across groups; all other parameters were freely estimated across groups.