NIH Public Access
Author Manuscript
Child Dev. Author manuscript; available in PMC 2016 January 01.

Published in final edited form as:
Child Dev. 2015 January ; 86(1): 159–175. doi:10.1111/cdev.12292.

Developmental Relations Between Vocabulary Knowledge and Reading Comprehension: A Latent Change Score Modeling Study

Jamie M. Quinn, **Richard K. Wagner**, **Yaacov Petscher**, and **Danielle Lopez**
Florida State University, Florida Center for Reading Research

Abstract

The present study followed a sample of first grade students (*N* = 316, mean age = 7.05 at first test) through fourth grade to evaluate dynamic developmental relations between vocabulary knowledge and reading comprehension. Using latent change score modeling, competing models were fit to the repeated measurements of vocabulary knowledge and reading comprehension to test for the presence of leading and lagging influences. Univariate models indicated growth in vocabulary knowledge and reading comprehension was determined by two parts: constant yearly change and change proportional to the previous level of the variable. Bivariate models indicated previous levels of vocabulary knowledge acted as leading indicators of reading comprehension growth, but the reverse relation was not found. Implications for theories of developmental relations between vocabulary and reading comprehension are discussed.

Prior to school entry and learning how to read, children learn vocabulary largely from social interactions with significant others in their environments (Phythian-Sence & Wagner, 2007). Rates of vocabulary development vary widely (Rowe, Raudenbush, & Goldin-Meadow, 2012), with some of this variability accounted for by differences in parent language and family background factors (Hoff, 2006). Although the meanings of some words are taught directly, inferring meanings of words from context appears to be an important component of vocabulary development (McKeown, 1985; Sternberg, 1987). Correlations between vocabulary knowledge and reading comprehension are substantial, ranging from .3 to .8 (Tannenbaum, Torgesen, & Wagner, 2006), with a tendency toward larger correlations with increasing development and reading experience (Torgesen, Wagner, Rashotte, Burgess, & Hecht, 1997).

Anderson and Freebody (1981) proposed three hypotheses that could account for correlations between vocabulary knowledge and reading comprehension. The first hypothesis, the instrumentalist hypothesis, posits a causal influence of vocabulary knowledge on reading comprehension. Simply put, the better one's knowledge of the meanings of the words in a passage, the better one's ability to comprehend the passage. The remaining two hypotheses posit non-causal relations between the two variables. The knowledge hypothesis states that vocabulary and reading comprehension are correlated

Corresponding author: quinn@psy.fsu.edu.

because they both related to the third variable of conceptual knowledge. The aptitude hypothesis states that vocabulary knowledge and reading comprehension are correlated because they both are related to the third variable of verbal aptitude.

Instrumental (Causal) Relations between Vocabulary and Reading Comprehension

Reading comprehension is supported by knowledge of words, including the precision of orthographic, phonological, and semantic representations (Perfetti & Hart, 2001; Verhoeven & van Leeuwe, 2008). Conversely, limited vocabulary knowledge or an inability to access vocabulary knowledge efficiently is believed to result in poor reading comprehension (Beck, Perfetti, & McKeown, 1982). In line with these views and with the instrumentalist hypothesis, vocabulary knowledge is one of the best-known predictors of reading comprehension (e.g., Beck & McKeown, 1991; Cain & Oakhill, 2011; McKeown, Beck, Omanson, & Perfetti, 1983; Wagner, Rashotte, Burgess, & Hecht, 1997). Specifically, receptive vocabulary breadth is a unique predictor above and beyond nonverbal IQ, decoding, and visual word recognition in the prediction of reading comprehension (Ouellette, 2006), and children with poor reading comprehension skills do not readily deduce meanings of novel words from context as well as skilled peers (Cain, Oakhill, & Lemmon, 2004).

The fact that vocabulary knowledge is a one of the most important predictors of reading comprehension is not sufficient proof of an instrumental relation, but the absence of a predictive relation would be problematical to an instrumental hypothesis. More direct support would come from studies that manipulate vocabulary knowledge and look for effects on reading comprehension. Although there is some support for an effect of training vocabulary knowledge on comprehension (e.g. Beck & McKeown, 1991, Cain & Oakhill, 2011, McKeown et al, 1983; Ouellette, 2006), and vocabulary training to improve reading comprehension was recommended by the National Reading Panel (NICHD, 2000), the results are mixed at best. Meta-analyses by Stahl and Fairbanks (1986) and Elleman, Lindo, Morphy, and Compton (2009) reported moderate effects of vocabulary training on experimenter-designed measures of comprehension that were created to be sensitive to the specific intervention under study, but little or no effects on broader, standardized measures of reading comprehension. Although it might be concluded that an absence of solid support for an effect of vocabulary training on reading comprehension dooms an instrumental hypothesis, instrumental relations can exist between variables that are not easily manipulated. For example, the annual tilting of the earth’s axis of rotation causes seasonal change, an instrumental relation that holds despite the fact that the angle of the earth’s rotation cannot be manipulated.

An alternative hypothesis concerning the relation between vocabulary and reading comprehension, not considered by Anderson and Freebody (1981), is that an instrumental relation may exist, but in the opposite direction, with reading comprehension having an instrumental effect on vocabulary knowledge (Wagner & Meros, 2010). Inference from context, a comprehension skill, is important for understanding text and is also considered a means of vocabulary learning (Cain, 2007; Nagy & Scott, 2000). For example, Cain and

Oakhill (2011) found that comprehension predicted later performance on a measure of receptive vocabulary (see also, Nation, Snowling, & Clarke, 2007). Poor reading comprehension or limited reading experience is believed to hamper vocabulary development (Cunningham & Stanovich, 1991; Nagy & Anderson, 1984; Nagy & Scott, 2000).

A second possibility not considered by Anderson and Freebody (1981) is that vocabulary knowledge plays an instrumental role in reading comprehension but the effect is mediated rather than direct (Wagner & Meros, 2010). For example, vocabulary knowledge is related to both phonological awareness and decoding (Lonigan, 2007). Phonological awareness refers to the ability to recognize and use the structure of oral language (Stanovich, 1992; Wagner & Torgesen, 1987) and decoding involves silently reading isolated words with speed and accuracy (Gough & Tunmer, 1986). Vocabulary knowledge may improve either phonological awareness or decoding, and both of these variables appear to be important for learning to read for comprehension (Lonigan, 2007; Lonigan, Burgess, & Anthony, 2000; Wagner, Torgesen, Laughon, Simmons, & Rashotte, 1993; Wagner and Meros, 2010).

Third Variable (Non Causal) Relations between Vocabulary Knowledge and Reading Comprehension

Turning to Anderson and Freebody's (1981) non-causal hypotheses about relations between vocabulary and reading comprehension, the knowledge hypothesis posits that vocabulary and reading comprehension are correlated because they both related to the third variable of conceptual knowledge. Their aptitude hypothesis posits that vocabulary and reading comprehension are correlated because they both are related to the third variable of verbal aptitude. On IQ tests, the vocabulary subtest is the best single subtest estimate of verbal IQ (Sattler, 2001), and measures of reading comprehension also are strongly correlated with verbal IQ (Woodcock, McGrew, & Mather, 2001, 2007).

Another example of a third-variable, aptitude hypothesis is provided by metalinguistic awareness, the awareness of and the ability to manipulate the sound structure of oral language (Nagy, 2007). Vocabulary knowledge is believed to be related to metalinguistic awareness in the form of morphological awareness, which is a form of metalinguistic awareness that refers to knowledge about word roots, prefixes, and suffixes and is believed to play a role in vocabulary development (Carlisle, 2007). Reading comprehension in turn is related to metalinguistic awareness in the form of metacognition, and more specifically comprehension monitoring, which is required for successful reading comprehension (Nagy, 2007).

Modeling Longitudinal Development of Vocabulary and Reading Comprehension

In addition to studies just discussed that used either simple correlations or training to examine relations between vocabulary knowledge and reading comprehension, a third approach is to model their longitudinal development. For example, Muter, Hulme, Snowling, and Stevenson (2004) reported that word identification, vocabulary, and linguistic skills at age 6 predicted comprehension skills in second grade. Oakhill, Cain, and Bryant

(2003) reported that verbal IQ, vocabulary, and inference skills predicted comprehension in grades three, four, and six. These two studies were simple prediction studies as opposed to causal models that controlled for autoregression. Controlling for autoregression refers to including the effect of previous state of a construct when predicting its future state along with other predictors of interest. This is important for attempting to make inferences about underlying causal influences from longitudinal data because failing to control for autoregression can result in spurious associations between variables (Gollob & Reichardt, 1987). For example, suppose vocabulary is measured at time one and reading comprehension is measured at time two. Time one vocabulary might predict time two reading comprehension only because time one vocabulary is correlated with unmeasured time one reading comprehension, and time one reading comprehension is the primary cause of time two reading comprehension.

Two studies that included autoregressor variables in their models also found vocabulary to predict subsequent reading comprehension. Third-grade vocabulary, decoding, and listening comprehension predicted fifth-grade reading comprehension after controlling for third-grade reading comprehension (de Jong & van der Leij, 2002). Verhoeven and van Leeuwe (2008) reported a strong influence of vocabulary on subsequent reading comprehension and a weak influence of reading comprehension on subsequent vocabulary in a study that included appropriate autoregressor variables.

Modeling Dynamic Longitudinal Influences with Latent Change Score Models

Until recently, longitudinal data useful for studying relations between vocabulary knowledge and reading comprehension have been modeled using one of two approaches. The first approach is cross-lagged regression or structural equation causal models using latent variables. The second approach is latent growth curve modeling of constructs represented by either single observed variables or latent variables with multiple indicators. Each of these approaches has strengths and weaknesses for fully exploring developmental relations among constructs.

An important strength of cross-lagged regression or SEM causal models is the facilitation of causal inference provided by time precedence. The developmental period is broken into discrete time intervals, thereby allowing one to determine whether construct A at time one predicts construct B at time 2, controlling for the autoregressor of time 1 construct B. A disadvantage of cross-lagged regression or SEM causal models is that because only covariance structures are modeled, actual development or growth is ignored. For example, identical model parameters and model fits statistics would be obtained regardless of whether performance on the measures grew, declined, or stayed the same over time.[1]

A major strength of latent growth-curve models is that growth is modeled explicitly with a developmental function. Unlike cross-lagged regression or SEM causal models that ignore means, latent growth curve models model all of the data including means as well as covariances. However, the developmental function applies to the entire developmental period, and because the developmental period is not broken into discrete time intervals, the absence of time precedence limits the ability to make causal inferences from the results of latent growth curve modeling.

Latent-change score models represent a way to model longitudinal data that combines the strengths of SEM causal models and latent growth curve models (McArdle, 2009). As is done with SEM causal models, the developmental period is divided into discrete time intervals and time-precedence is available to be used for making causal inferences. As is true for latent growth curve models, development is modeled explicitly by analyzing both mean structures and covariance structures. Further, with the specification of a latent change score as a higher-order factor, it allows for the test of individual differences in change though rejecting or not rejecting the null hypothesis of zero variance. The key contribution of latent change scores models, however, is that they are useful for modeling dynamic relations between constructs as they develop over time (Ferrer & McArdle, 2010). Dynamic relations between constructs are described in terms of leading and lagging indicators: A construct is considered a leading indicator if change in this construct "leads" change in another construct. The second construct is therefore considered a "lagging indicator" of the first: its development is coupled with but lagging behind that of the first construct. Coupling parameters in latent change score models are used to represent these dynamics by determining whether performance on one construct can account for subsequent change in performance on a second construct. Latent change score models thus provide an important window into possible underlying causal influences between constructs that change over time because of growth and development but are not easily manipulated via training studies.

Recently, a few studies on dynamic longitudinal relations between general cognitive abilities and reading abilities have been published. Ferrer et al. (2007) investigated dynamic longitudinal relations between a reading composite that included the letter-word identification, decoding, and comprehension subtests of the Woodcock-Johnson R (WJ-R) and the subscales of the Wechsler Intelligence Scales for Children (WISC-R). The sample was from the Connecticut Longitudinal Study, and included 445 students who were followed from 1st through 12th grade. Of most interest to the present study, bidirectional coupling influences were reported between a reading composite of reading and IQ. In other words, relative performance at a given time in verbal IQ signaled subsequent change in reading comprehension, and conversely, relative performance at a given time in reading comprehension signaled subsequent change in IQ. This coupling was stronger for performance IQ than for verbal IQ, and for younger children than for older children. This result also held when looking specifically at coupling between verbal IQ and the passage comprehension subtest, which are the two measures most similar to those used in the present study. Ferrer et al. (2010) reported a second analysis on a subset of the same Connecticut Longitudinal Study sample in which the sample was sorted into three groups: typical readers, compensated readers, and persistently poor readers. Their previously reported finding of coupling between reading and IQ held only for the typical readers. No coupling was found for either the compensated readers or the persistently poor readers.

In an attempt to replicate the findings of Ferrer et al. (2007, 2010), Reynolds and Turek (2012) investigated dynamic relations between verbal-comprehension knowledge (Gc) and reading comprehension at 3rd, 5th, and 9th grades. Their measure of verbal-comprehension knowledge was picture vocabulary from the Woodcock Johnson-Revised (Woodcock & Johnson, 1989). Their results conflicted with those of Ferrer et al. (2007, 2010) in that only unidirectional coupling from vocabulary to reading comprehension was found. There was no corresponding influence of reading comprehension on subsequent vocabulary growth. Reynolds and Turek (2012) speculated on two possible causes of the different results between their study and Ferrer et al. (2007, 2010). First, the Connecticut Longitudinal Study used by Ferrer et al. began in first grade, whereas the Reynolds and Turek study did not begin until third grade. There is evidence that the nature of reading comprehension changes from first through third grade, depending less on decoding and depending more on listening comprehension (Kim, Wagner, & Lopez, 2012). Second, the studies used different measures for both reading and aptitude. We also would note that both of the Ferrer et al. studies used the same Connecticut Longitudinal Study dataset. With Reynolds and Turek using a second dataset, only two datasets have been modeled. There clearly is a need to apply similar models to additional datasets.

The Present Study

The purpose of the present study was to investigate potential developmental coupling of vocabulary and reading comprehension using latent change score modeling. Our focus was to examine the co-development of reading comprehension and vocabulary knowledge as a construct, as opposed to IQ or crystallized ability more generally (e.g., Ferrer et al., 2007, 2010; Reynolds & Turek, 2012), because of the hypothesized special relation between acquisition of vocabulary knowledge and reading comprehension (Anderson & Freebody, 1981; Wagner & Meros, 2010). We first modeled growth of each construct separately by comparing the fit of three models to the vocabulary knowledge and reading comprehension data. The first model was a constant-change model that posited that growth was linear. This model was equivalent to a latent growth curve model with a linear growth function. The second model was a proportional change model that posited that growth was a function of previous level of performance. The third model was a dual-change model that incorporated both linear and proportional change components. Both the constant-change model and the proportional change model were nested within the dual-change model, allowing nested model testing to be used to evaluate the models.

Once growth was modeled separately for vocabulary knowledge and reading comprehension, we combined the models into a bivariate model that tested for coupled relations in the co-development of vocabulary knowledge and reading comprehension. Specifically, we sought to test four competing hypotheses about these developmental relations that could be implemented as alternative latent change score models:

1. Correlated but uncoupled development

According to this hypothesis, development of vocabulary is potentially correlated but not coupled with that of reading comprehension. In other words, children who grow faster in vocabulary also grow faster in reading comprehension, but there is no temporal coupling of development, in that level of performance in one construct does not account for subsequent year to year change in the other construct. Anderson and Freebody's (1981) aptitude and knowledge hypotheses, in which correlations between vocabulary and reading comprehension are attributed to their joint correlation with third variables, are examples of correlated but uncoupled developmental hypotheses. The latent change score model that corresponds to the correlated but uncoupled developmental hypothesis is one in which (a) slope (i.e., growth) in vocabulary is allowed to be correlated with slope in reading comprehension, (b) the intercept for vocabulary is allowed to be correlated with the intercept for reading comprehension, but (c) no coupling is allowed.

2. Unidirectional coupling from vocabulary knowledge to reading comprehension

According to this hypothesis, subsequent changes in reading comprehension are accounted for in part by current levels of vocabulary knowledge. Anderson and Freebody's (1981) instrumental hypothesis is an example of a unidirectional coupling from vocabulary to reading comprehension hypothesis. The corresponding latent change score model is one in which (a) vocabulary slope and intercept are allowed to be correlated with reading comprehension slope and intercept, (b) coupling is allowed from vocabulary to change in reading comprehension, but (c) no coupling is allowed from reading comprehension to change in vocabulary.

3. Unidirectional coupling from reading comprehension to vocabulary

According to this hypothesis, subsequent changes in vocabulary are accounted for in part by current levels of reading comprehension. The idea is that text provides an opportunity to learn new vocabulary, and more skilled reading comprehension facilitates this process (Nagy & Anderson, 1984). The corresponding latent change score model is one in which (a) vocabulary slope and intercept are allowed to be correlated with reading comprehension slope and intercept, (b) coupling is allowed from reading comprehension to change in vocabulary, but (c) no coupling is allowed from vocabulary to change in reading comprehension.

4. Bidirectional coupling model

This hypothesis represents a combination of the unidirectional coupling from vocabulary to reading comprehension and the unidirectional coupling from reading comprehension hypotheses. The corresponding latent change score model allows (a) correlations between vocabulary knowledge and reading comprehension slopes and intercepts, (b) coupling from vocabulary knowledge to reading comprehension, and (c) coupling from reading comprehension to vocabulary knowledge.

Models 1 through 3 are nested in model 4, and model 1 is nested within models 2 and 3, which enabled us to use chi-square difference testing to determine whether there were statistically significant differences in the fit of the models to the data.

Method

Participants

The participants in this study were from the Florida Longitudinal Study, a longitudinal study of the co-development of language and literacy from first through fourth grades. A total of 316 first grade children from schools in the Leon County School District began with the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

study in 2007. The sample demographics were 60% White, 25% Black, 4% Hispanic, 4% Asian, and 7% other. The sample was 51.6% male. The sample was primarily English speaking. Participants' mean age at the initial date of testing was 7.05 with a range from 6.14 to 8.80 years of age. At final testing, 219 children remained in the study, which represented an annual rate of attrition of 10 percent. Their mean age was 9.85 with a range from 9.10 to 11.09 years of age. Missing data were handled using full-information maximum likelihood in Mplus 7 during model estimation.

Measures

The following measures of vocabulary knowledge and reading comprehension were individually administered.

Stanford-Binet Intelligence Scales V: Vocabulary Subtest—The Vocabulary subtest of the Stanford-Binet Intelligence Scales was used as a measure of breadth of expressive vocabulary. Internal consistency averages 0.87, and test-retest reliability averages 0.75. Found to be highly correlated with the verbal subtests of the Wechsler Intelligence Scales for Children (WISC-R), Wechsler Preschool and Primary Scale of Intelligence (WPPSI), and Wechsler Adult Intelligence Scale (WAIS-R; range 0.72 – 0.86) (Thorndike, Hagen, & Sattler, 1986).

Wechsler Abbreviated Scales of Intelligence: Vocabulary Subtest—The Wechsler Abbreviated Scales of Intelligence (WASI) Vocabulary subtest is an individually administered test of breadth of expressive vocabulary. Test-retest reliability coefficients average 0.85; inter-rater reliability coefficients average 0.98. The WASI has been found to highly correlate with the Wechsler Adult Intelligence Scale (WAIS) and the Wechsler Intelligence Scales for Children (WISC-III; range 0.72 – 0.88) and moderately correlate with the WIAT subtests (0.57 – 0.66) (The Psychological Corporation, 1999).

Woodcock Reading Mastery Test – Revised/Normative Update: Passage comprehension—The Woodcock Reading Mastery Test (WRMT) Passage Comprehension subtest measures reading comprehension using a cloze procedure in which the child is asked to read a short passage (usually 2 to 3 sentences long) and identify the missing key words. The passage comprehension subtest is part of the Reading Comprehension Cluster of the WRMT. Internal consistency reliability is 0.91, split-half reliability averages 0.97 (range .86 – .99). Highly correlated with WJ reading tests, Woodcock Reading Achievement Test (WRAT) and Iowa Tests of Basic Skills (range 0.71 – 0.92) (Woodcock, 1987).

Woodcock-Johnson III Tests of Achievement: Passage Comprehension—The Woodcock-Johnson III Passage Comprehension subtest (WJPC) also assesses reading comprehension using a cloze procedure with children reading short passages and identifying missing key words that makes sense in the context of that passage. Median split-half reliability is 0.88. The reading comprehension cluster correlates moderately to highly with the Wechsler Individual Achievement Test (WIAT; .70–.79) and the Kaufman Test of Educational Achievement (KTEA; .62–.81) (McGrew & Woodcock, 2001).

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Procedures

Trained research assistants individually administered all measures annually. The subtests were administered as part of the larger longitudinal study. A fixed order of presentation was used, with multiple measures of the same constructs administered during different testing sessions to eliminate time sampling error from the latent variables.

Results

Descriptive Statistics

Table 1 contains the sample sizes, means, standard deviations, and correlations of the raw scores from grades 1 through 4. Standardized, age-based score averages and ranges are also included for reference purposes. For WJPC and WRMT, scores are standardized with a mean of 100 and standard deviation of 15, and scores between 90 and 110 are considered average for similarly aged peers (McGrew & Woodcock, 2001; Woodcock, 1987). WASI and SB were standardized using *t*-scores with a mean of 50 and standard deviation of 10. There is a pattern of increasing performance over time, as is seen in Figure 1, with all four measures plotted over time. For the WJPC and the WRMT data, there is increasing performance along with decreasing variance. The Vocabulary subtest from the Stanford-Binet shows relatively stable variance over time, and that from the WASI shows increasing variability coupled with increasing performance. There are moderate to relatively high correlations between the four measures over time. The correlations are based on Full Information Maximum Likelihood estimation (FIML; e.g. McArdle, 1994), an estimation method that allows for the examination of sample descriptive statistics as if all members of the sample were present at all measurement occasions. We compared these correlations based on FIML as well as the means and standard deviations with the typical correlations, means, and standard deviations based on pair-wise deletions and with all available data and found them to be nearly identical, which suggests that these data meet minimum requirements for “missing at random” (MAR; Little, 1995).

The scaling of the latent change score is an important consideration. Without meaningful scaling of scores over time, the resulting change scores are not interpretable. Observed scores were converted to z-scores using the means and standard deviations from the first time point for proper scaling. Consequently, the unit of the latent change score models is standardized unit change relative to the variability observed at the first time point. These developmentally scaled z-scores were each regressed on to a latent variable representing the constructs of interest: vocabulary knowledge and reading comprehension (see Figure 1).

Model fit was assessed using the chi-squared (χ^2) test of model fit statistic, the root mean squared error of approximation (RMSEA), the comparative fit index (CFI), and the Tucker-Lewis Index (TLI). Comparison of nested models was made using a chi-squared difference test where applicable and non-nested models were compared using the Bayesian information criteria (BIC; Schwarz, 1978), where lower values indicated better fit.

Univariate Latent Change Score Models

A series of competing models were fit to the data separately for reading comprehension and vocabulary knowledge. Beginning with reading comprehension, a univariate dual change score model fit the data moderately well, χ^2 (35) = 95.81, p < .001, CFI = .971, TLI = .977, RMSEA = .074 (95% confidence interval [CI]: .057–.092), BIC = 3283.69. Modification indices showed fit could be improved by relaxing some of the constraints, but none of the modifications were theoretically sound. Removing the constant change portion of the model and only estimating proportional change resulted in severe degradation of model fit, $\Delta\chi^2$ (38) = 1851.86, p < .001, CFI = .140, TLI = .366, RMSEA = .389 (95% CI: .374 – .404), BIC = 5022.48; χ^2 (3) = 1756.05, p < .001. Similarly, removing the proportional change portion of the model and only estimating constant change also resulted in severe degradation of model fit, χ^2 (36) = 1869.72, p < .001, CFI = .130, TLI = .324, RMSEA = .401, BIC = 5051.85; $\Delta\chi^2$ (1) = 1773.91, p < .001. This indicated that adequate modeling of growth required both constant change and proportional change parameters. Growth was characterized by a constant amount of change and a decreasing proportional change that attenuated growth over time.

Next, these models were fit to vocabulary knowledge. A dual change model fit the data well, χ^2 (35) = 54.72, p < .001, CFI = .985, TLI = .988, RMSEA = .042 (95% CI: .018 – .063), BIC = 5027.76. Removing the constant change portion of the model resulted in degradation of model fit, χ^2 (38) = 794.38, p < .001, CFI = .422, TLI = .574, RMSEA = .251 (95% CI: .236 –.266), BIC = 5750.15; $\Delta\chi^2$ (3) = 739.66, p < .001. Removing the proportional change portion of the dual change model also resulted in a small degradation of fit, χ^2 (36) = 59.82, p < .001, CFI = .982, TLI = .986, RMSEA = .046, BIC = 5027.10.15; $\Delta\chi^2$ (1) = 5.1, p < .02. Results of the chi-square difference test show removing this parameter has a significant impact on model fit; the other fit statistics change only slightly. However, the proportional change parameter was left in the model to help explain univariate growth (the parameter was significant in the dual-change model). Therefore, as was the case for reading comprehension, growth was characterized by a combination of a constant amount of change and a decreasing proportional amount of change that attenuated rate of growth over time.

Parameter estimates from the best-fitting univariate dual-change score models of vocabulary knowledge development and of reading comprehension development are presented as path models in Figures 2 and 3. The mean intercepts, or average initial scores at the first time point, were not significantly different from zero (μ_0 = −0.010 for vocabulary knowledge and −0.004 for reading comprehension), as expected due to the conversion of raw scores to z-scores. There was significant variation in the initial means (μ_0), indicating substantial individual differences in the starting values for both vocabulary knowledge and reading comprehension. There was significant growth (μ_1) in vocabulary knowledge (h_1 = 0.790) and reading comprehension (g_1 = 1.044). The units of these mean slopes are interpretable as first-grade standard deviation units. There also was significant variation in amount of growth for both constructs (σ_{h1} = 0.042; σ_{g1} = 0.065). The proportional change parameter was significant and negative for both vocabulary knowledge (β_v = −0.081) and reading comprehension (β_c = −0.416), indicating an overall slowing of growth over time. The positive correlations between individual differences in intercept and slope for both vocabulary knowledge and reading comprehension ($\sigma_{0,1}$ = 0.293; $\sigma_{0,1}$ = 0.621, respectively) are indicative of fan-spread growth: Higher scores at the first time point were associated with higher rates of growth. This effect was more evident for growth in reading comprehension compared to growth in vocabulary knowledge.

Bivariate Dual Change Score Models

Having successfully modeled growth in vocabulary knowledge and reading comprehension separately, it was then possible to explore potential dynamic relations between their co-development by modeling them simultaneously using a bivariate dual change score model. The critical additional parameters that were estimated in the bivariate dual change score models were the coupling parameters. The coupling parameters were indicators of the extent to which individual differences in level of performance for one construct could account for individual differences in subsequent growth for the other construct. Covariances were also estimated between reading comprehension and vocabulary knowledge slopes and intercepts.

A full, bidirectional-coupling model fit the data well (χ^2 [124] = 199.71, p < .001, CFI = .980, TLI = .980, RMSEA = .044 [95% CI: .032 – .055], BIC = 8062.76). Using identical logic to that used for univariate model comparison, we next tested two models that were nested in the bidirectional-coupling model. A vocabulary knowledge to reading comprehension coupling only model eliminated the coupling paths from reading comprehension to vocabulary knowledge from the full bidirectional model. This did not result in a significant decrement in model fit, $\Delta\chi^2[1] = 2.4$, p = .121, compared to the bidirectional model. A reading comprehension to vocabulary knowledge coupling only model eliminated the coupling paths from vocabulary to reading comprehension from the full bidirectional model. This resulted in a significant degradation in fit, $\Delta\chi^2[1] = 7.85$, p < .01. Lastly, eliminating all cross-construct coupling pathways from the bidirectional model leaving only a correlated growth model also resulted in significant degradation in fit, $\Delta\chi^2[2] = 11.61$, p < .01.

In summary, the results of the chi-square difference testing supported the vocabulary knowledge to reading comprehension coupling only model. This model, which is presented in Figure 4, provided a good fit to the data, with χ^2 [125] = 202.11, p < .001, CFI = .979, TLI = .980, RMSEA = .044 (95% CI: .033 – .055), BIC = 8059.41. The positive correlations between slope and intercept within construct (vocabulary: r = 0.317; reading comprehension: r = 0.656) reflect the fan-spread growth that was noted in the univariate models. The positive correlations between intercept of vocabulary and slope of reading comprehension (r = 0.496), between intercept of reading comprehension and slope of vocabulary (r = 0.371), and between intercept of vocabulary and intercept of reading comprehension (r = 0.663), reflect the fact that initial levels and growth in vocabulary knowledge and reading comprehension are positively correlated, such that higher initial level in one construct leads to larger growth in the other construct. To fulfill measurement invariance criteria, within construct errors (e1–e4) were constrained to equality over time. The larger values for the two measures of vocabulary knowledge (WASI = 0.330; SB = 0.471) show that the vocabulary latent variable did not account for as much variance in its indicators compared to the variance in the two reading comprehension measures (WJPC =

0.179; WRMT = 0.111) accounted for by the reading comprehension latent variable. Further, these within construct error terms were allowed to correlate within but not across time to allow for proper model estimation. Only two cross-construct residual correlations were significant, SB with WJPC ($r = 0.075$, $p < .041$) and WASI with WRMT ($r = 0.107$, $p < .007$), and the other correlations failed to reach significance: WJPC with WASI ($r = −.015$, $p = .696$) and WRMT with SB ($r = 0.061$, $p = .108$). Of most interest, however, is the significant and positive coupling parameter from vocabulary to reading comprehension of .184 standardized units, which indicates that annual growth in reading comprehension was accounted for in part by level of vocabulary knowledge. Specifically, a student whose level of vocabulary was 1 standard deviation above that of an average student would grow .184 standard deviations faster in reading comprehension over a year. Figure 5 presents the estimated growth trajectories using the model parameter estimates from this bivariate latent change model.

Discussion

Latent change score modeling was used to investigate dynamic relations between the development of vocabulary knowledge and reading comprehension. First, development of vocabulary knowledge and reading comprehension were modeled separately. The purpose was to identify a best-fitting univariate model of growth for each construct prior to investigating potential coupling relations. The results indicated that development of both vocabulary knowledge and reading comprehension was adequately described by dual-change models in which growth reflected both positive constant change and negative proportional change. In other words, growth occurred every year but the rate of this growth diminished over time. For both constructs, a positive correlation between slope and intercept indicated that the pattern of growth was fan-spread, with higher initial performers growing faster than lower initial performers.

Next, dynamic relations in the co-development of vocabulary knowledge and reading comprehension were investigated by comparing alternative bivariate latent change score models. A nested model comparison supported the unidirectional coupling from vocabulary to reading comprehension model over the correlated but uncoupled growth, unidirectional coupling from reading comprehension to vocabulary, or bidirectional coupling models. The present results support Anderson and Freebody’s (1981) instrumentalist hypothesis that vocabulary knowledge has a causal influence on reading comprehension. More specifically, the results support a developmental generalization of Anderson and Freebody’s instrumentalist hypothesis in that they proposed their hypothesis as an explanation of observed correlations between vocabulary knowledge and reading comprehension at a given point of time. They did not consider dynamic developmental relations between vocabulary knowledge and reading comprehension over time.

Turning to previous latent change score modeling studies of the development of reading comprehension and aptitude, our results are consistent with those of Reynolds and Turek (2012), who also found one-way coupling from vocabulary knowledge to reading comprehension as opposed to the bi-directional coupling between verbal aptitude and reading comprehension reported by Ferrer et al. (2007, 2010). In discussing possible reasons

NIH-PA Author Manuscript

NIH-PA Author Manuscript

for the differences in results between the Reynolds and Turek and Ferrer et al. studies, Reynolds and Turek suggested that a difference in starting points might have been responsible for differences in results. The Ferrer et al. studies began in first grade, but Reynolds and Turek modeled growth beginning in third grade. The results of the present study cast doubt on this explanation because the results of our study were similar to those of Reynolds and Turek despite the fact that our study began in first grade, similarly to the Ferrer et al. studies. A comparison of the design of the three studies suggests another possible explanation for the difference in results between the Ferrer et al. studies and both the Reynolds and Turek and present studies, namely that the results may be different for vocabulary knowledge compared to cognitive ability more broadly defined. Although the focus of Reynolds and Turek's study was crystallized intelligence (Gc), they used Picture Vocabulary from the Woodcock Johnson as their measure of Gc. We also used measures of vocabulary in the present study. In contrast, Ferrer et al. used the broader Verbal and Performance scales from the WISC-R as their measures of cognitive ability, and reported even greater coupling from reading comprehension to cognitive ability for the Performance Scale than for the Verbal Scale. It will be important to carry out future studies with a broader array of vocabulary measures that target depth as well as breadth and go beyond simple definitional knowledge.

The results of the present study need to be considered in the context of several constraints and limitations. First, we studied development as it occurred naturally given the instructional practices that were used in the schools at the time of our study. Instructional practices influence academic development in early elementary students (Cameron, Connor, & Morrison, 2005; Rimm-Kaufman, La Paro, Downer, & Pianta, 2005), and may explain as much as one-third of the variation in student achievement across a school year (Creemers & Reezigt 1996). If we assume that current instructional practices either target or affect reading more than vocabulary, the likely effect on the results would be to strengthen the reading to vocabulary coupling. Because we found the opposite pattern of results, with vocabulary influencing reading, classroom instructional practices are not a likely explanation of our results. It is certainly possible that other approaches to teaching reading and facilitating language development might have led to different outcomes.

Further, although these students all come from the same school district within the state of Florida, it is possible that differences in curriculum implementation across schools within the district could explain some of the variance in growth and initial levels. Data on classroom instructional practices were not available. Second, although it is appropriate to make causal inferences from the results of modeling longitudinal data (Pearl, 2012), caution is warranted by the fact that this study did not experimentally manipulate vocabulary knowledge. Experimental or intervention studies represent a more direct test of causal relations between constructs, and converging results from intervention studies and latent change score models would be reassuring. It is important to consider, however, the reasons why results from latent change score modeling studies of development and intervention studies might not converge. For example, results from latent change score modeling of longitudinal data might support a causal inference that is not supported by intervention studies if it is not possible to substantially affect the level of a construct through a typical intervention but change and individual differences in change emerge in the course of normal development. On the other hand, a causal inference might be supported by intervention studies and not by latent change score modeling of development if intervention results in a substantial change in the level of a construct relative to what typically occurs with development.

Elleman et al.'s (2009) meta-analysis of the effects of vocabulary intervention on reading comprehension is particularly relevant in the present context. Based on a meta-analysis of 37 studies, they found that vocabulary intervention had a significant effect size of .50 for custom reading comprehension measures that were designed to be sensitive to the vocabulary invention, but a non-significant effect size of .10 for standardized measures of reading comprehension. Given that the present study used standardized measures of both vocabulary and reading comprehension, the effect size for standardized measures of reading comprehension provides the most apt comparison. Because the variables in the present study were standardized based on first-grade means and standard deviations, the coupling parameter of .184 can be interpreted as the difference in annual growth in reading comprehension attributable to a standard deviation difference in vocabulary knowledge. It is interesting that the magnitude of the significant coupling parameter is not that much different than the magnitude of the non-significant effect size ($d = .10$) found in the meta-analysis. One interesting difference is that the coupling parameter represents the effect of change in vocabulary over an entire year, whereas the typical vocabulary intervention included in Elleman et al.'s (2009) meta-analysis was well under a year in duration. In the present study, the coupling effect of .184 represented the effect of vocabulary knowledge on growth in reading comprehension each year. The cumulative effect across three years of the longitudinal study was .54. It would be interesting to know whether the effects of vocabulary intervention on reading comprehension would also accumulate across years if an intervention were carried out over multiple years as opposed to less than a single year in a typical intervention study. Small effects can have large results if the effects accumulate, and educational interventions are examples of effects that are likely to accumulate (Abelson, 1985).

Evidence for a coupling from level of reading comprehension to growth in vocabulary knowledge was not found in the present study. The results of our nested model testing supported the vocabulary knowledge to reading comprehension coupling only model. Relatedly, when the reading comprehension to vocabulary knowledge coupling parameter was estimated in the initial full, bidirectional coupling model it was not significant. However, additional studies—particularly large N studies—using other measures of vocabulary knowledge and reading comprehension are needed before dismissing an effect of reading comprehension on subsequent growth in vocabulary knowledge. In the present study, the magnitude of the reading comprehension to vocabulary knowledge coupling parameter estimated in the full, bidirectional coupling model was .154 ($p = 0.099$), a value not that much different than the significant vocabulary knowledge to reading comprehension coupling parameter of .158 ($p < .005$). It also is important to remember that our results are dependent on the measures of vocabulary knowledge and reading comprehension that we used. If our measures of vocabulary knowledge were less sensitive to change than were our measures of reading comprehension, it might be harder to find evidence of a coupling from reading comprehension to vocabulary knowledge than from vocabulary knowledge to reading comprehension. Similarly, the usage of only cloze task format reading comprehension tests has been criticized. These criticisms include that they are heavily dependent on phonological awareness and spelling (Mehta, Foorman, Branum-Martin, & Taylor, 2005) and word reading skills (Francis et al., 2006; Keenan, Betjemann, & Olson, 2008). The cloze format tests used in this study were the only tasks available in our longitudinal study. Future studies should include a variety of task formats to have a completely representative latent factor for reading comprehension.

In conclusion, the results of the present study support the idea that growth in reading comprehension depends in part on vocabulary knowledge. More broadly, latent change score modeling of the co-development of related constructs represents an important new tool for advancing our understanding of the complexities in child development.

Acknowledgments

Support for this research was provided by Grant P50 HD052120 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development and a Predoctoral Interdisciplinary Fellowship, funded by the Institute of Education Sciences, U.S. Department of Education (R305B04074 & R305B090021).

References

Abelson RP. A variance explanation paradox: When a little is a lot. Psychological Bulletin. 1985; 97:129–133.10.1037/0033-2909.97.1.129

Anderson, R.; Freebody, P. Vocabulary knowledge. In: Guthrie, JT., editor. Comprehension and teaching: Research reviews. Newark, DE: International Reading Association; 1981. p. 77-117.

Beck, I.; McKeown, M. Conditions of vocabulary acquisition. In: Barr, R.; Kamil, ML.; Mosenthal, PB.; Pearson, PD., editors. Handbook of reading research. New York: Longman; 1991. p. 789-814.

Beck I, Perfetti C, McKeown M. Effects of long-term vocabulary instruction on lexical access and reading comprehension. Journal of Educational Psychology. 1982; 74:506–521.10.1037/0022-0663.74.4.506

Cain K. Deriving word meanings from context: Does explanation facilitate contextual analysis? Journal of Research in Reading. 2007; 30:347–359.10.1111/j.1467-9817.2007.00336.x

Cain K, Oakhill J. Matthew effects in young readers: Reading comprehension and reading experience aid vocabulary development. Journal of Learning Disabilities. 2011; 44:431–443.10.1177/0022219411410042 [PubMed: 21772058]

Cameron CE, Connor CM, Morrison FJ. Effects of variation in teacher organization on classroom functioning. Journal of School Psychology. 2005; 43:61–85.10.1016/j.jsp.2004.12.002

Carlisle, JF. Fostering morphological processing, vocabulary development, and reading comprehension. In: Wagner, R.; Muse, A.; Tannenbaum, K., editors. Vocabulary acquisition: Implications for reading comprehension. New York: Guilford Press; 2007. p. 78-103.

Creemers BPM, Reezigt GJ. School level conditions affecting the effectiveness of instruction. School Effectiveness and School Improvement. 1996; 7:197–228.10.1080/0924345960070301

Cunningham A, Stanovich K. Tracking the unique effects of print exposure in children—association with vocabulary, general knowledge, and spelling. Journal of Educational Psychology. 1991; 83:264–274.10.1037/0022-0663.83.2.264

de Jong P, van der Leij A. Effects of phonological abilities and linguistic comprehension on the development of reading. Scientific Studies of Reading. 2002; 6:51–77.10.1207/S1532799XSSR0601_03

Elleman AM, Lindo EJ, Morphy P, Compton DL. The impact of vocabulary instruction on passage level comprehension of school age children: A meta-analysis. Journal of Research on Educational Effectiveness. 2009; 2:1–44.10.1080/19345740802539200

Ferrer E, McArdle JJ, Shaywitz BA, Holahan JM, Marchione K, Shaywitz SE. Longitudinal models of developmental dynamics between reading and cognition from childhood to adolescence. Developmental Psychology. 2007; 43:1460–1473.10.1037/0012-1649.43.6.1460 [PubMed: 18020824]

Ferrer E, McArdle JJ. Longitudinal modeling of developmental changes in psychological research. Current Directions in Psychological Science. 2010; 19:149–154.10.1177/0963721410370300

Ferrer E, Shaywitz BA, Holahan JM, Marchione K, Shaywitz SE. Uncoupling of reading and IQ over time: Empirical evidence for a definition of dyslexia. Psychological Science. 2010; 21:93–101.10.1177/0956797609354084 [PubMed: 20424029]

Francis DJ, Snow CE, August D, Carlson CD, Miller J, Iglesias A. Measures of reading comprehension: a latent variable analysis of the diagnostic assessment of reading comprehension. Scientific Studies of Reading. 2006; 10:301–322.10.1207/s1532799xssr1003_6

Gollob HF, Reichardt CS. Taking account of time lags in causal models. Child Development. 1987; 58:80–92.10.2307/1130293 [PubMed: 3816351]

Gough PB, Tunmer WE. Decoding, reading, and reading disability. Remedial and Special Education. 1986; 7(1):6–10.10.1177/074193258600700104

Hoff E. How social contexts support and shape language development. Developmental Review. 2006; 26:55–88.10.1016/j.dr.2005.11.002

Keenan JM, Betjemann RS, Olson RK. Reading comprehension tests vary in the skills they assess: differential dependence on decoding and oral comprehension. Scientific Studies of Reading. 2008; 12:281–300.10.1080/10888430802132279

Kim YS, Wagner RK, Lopez D. Developmental relations between reading fluency and reading comprehension: A longitudinal study from Grade 1 to Grade 2. Journal of Experimental Child Psychology. 2012; 113:93–111.10.1016/j.jecp.2012.03.002 [PubMed: 22726256]

Little R. Modeling the dropout mechanism in repeated-measures studies. Journal of the American Statistical Association. 1995; 90:1112–1121.10.1080/01621459.1995.10476615

Lonigan, CJ. Vocabulary development and the development of phonological awareness skills in preschool children. In: Wagner, R.; Muse, A.; Tannenbaum, K., editors. Vocabulary acquisition: Implications for reading comprehension. New York: Guilford Press; 2007. p. 15-31.

Lonigan CJ, Burgess SR, Anthony JL. Development of emergent literacy and early reading skills in preschool children: Evidence from a latent variable longitudinal study. Developmental Psychology. 2000; 36:596–613.10.1037/0012-1649.36.5.596 [PubMed: 10976600]

McArdle JJ. Structural factor analysis experiments with incomplete data. Multivariate Behavioral Research. 1994; 29:409–454.10.1207/s15327906mbr2904_5

McArdle JJ. Latent variable modeling of longitudinal data. Annual Review of Psychology. 2009; 60:577–605.10.1146/annurev.psych.60.110707.163612

McGrew, KS.; Woodcock, RW. Woodcock-Johnson III. Itasca, IL: Riverside Publishing; 2001. Technical Manual.

McKeown MG. The acquisition of word meaning from context by children of high and low ability. Reading Research Quarterly. 1985; 20:484–496.10.2307/747855

McKeown MG, Beck IL, Omanson RC, Perfetti CA. The effects of long-term vocabulary instruction on reading comprehension: a replication. Journal of Reading Behavior. 1983; 15:3–18.10.1080/10862968309547474

Mehta PD, Foorman BR, Branum-Martin L, Taylor WP. Literacy as a unidimensional multilevel construct: validation, sources of influence, and implications in a longitudinal study in grades 1 to 4. Scientific Studies of Reading. 2005; 9:85–116.10.1207/s1532799xssr0902_1

Muter V, Hulme C, Snowling MJ, Stevenson J. Phonemes, rimes, and language skills as foundations of early reading development: Evidence from a longitudinal study. Developmental Psychology. 2004; 40:663–681.10.1037/0012-1649.40.5.665

Nagy, WE. Metalinguistic awareness and the vocabulary-comprehension connection. In: Wagner, R.; Muse, A.; Tannenbaum, K., editors. Vocabulary acquisition: Implications for reading comprehension. New York: Guilford Press; 2007. p. 53-77.

Nagy WE, Anderson RC. How many words are there in printed school English? Reading Research Quarterly. 1984; 19:304–330.10.2307/747823

Nagy, WE.; Scott, JA. Vocabulary processes. In: Kamil, ML.; Mosenthal, P.; Pearson, D.; Barr, R., editors. Handbook of Reading Research. Vol. 3. Mahwah, NJ: Lawrence Eribaum; 2000. p. 269-283.

Nation K, Snowling MJ, Clarke P. Dissecting the relationship between language skills and learning to read: Semantic and phonological contributions to new vocabulary learning in children with poor reading comprehension. Advances in Speech Language Pathology. 2007; 9:131–139.10.1080/14417040601145166

National Institute of Child Health and Human Development. NIH Publication No. 00-4769. Washington, DC: U.S. Government Printing Office; 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.

Oakhill J, Cain K, Bryant PE. The dissociation of word reading and text comprehension: Evidence from component skills. Language and Cognitive Processes. 2003; 18:443–468.10.1080/01690960344000008

Ouellette G. What's meaning got to do with it: the role of vocabulary in word reading and reading comprehension. Journal of Educational Psychology. 2006; 98:554–566.10.1037/0022-0663.98.3.554

Pearl, J. The causal foundations of structural equation modeling. In: Hoyle, R., editor. Handbook of structural equation modeling. New York: Guildford Press; 2012. p. 68-91.

Perfetti, CA.; Hart, L. The lexical bases of comprehension skill. In: Gorfien, D., editor. On the consequences of meaning selection. Washington, DC: American Psychological Association; 2001. p. 67-86.

Phythian-Sence; Wagner, RK. Vocabulary acquisition: A primer. In: Wagner, R.; Muse, A.; Tannenbaum, K., editors. Vocabulary acquisition: Implications for reading comprehension. New York: Guildford Press; 2007. p. 1-14.

Psychological Corporation. Wechsler Abbreviated Scale of Intelligence (WASI) manual. San Antonio, TX: Author; 1999.

Reynolds MR, Turek JJ. 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. 2012; 50:841–863.10.1016/j.jsp.2012.07.002 [PubMed: 23245503]

Rimm-Kaufman SE, La Paro KM, Downer JT, Pianta RC. The contribution of classroom setting and quality of instruction to children's behavior in kindergarten classrooms. The Elementary School Journal. 2005; 105:377–394.10.1086/429948

Rowe ML, Raudenbush SW, Goldin-Meadow S. The pace of vocabulary growth helps predict later vocabulary skill. Child Development. 2012; 83:508–525.10.1111/j.1467-8624.2011.01710.x [PubMed: 22235920]

Schwarz G. Estimating the dimension of a model. The Annals of Statistics. 1978; 6:461–464.

Sattler, J. Assessment of children: Cognitive applications. 4. La Mesa, CA: Jerome M Sattler Publisher; 2001.

Stahl S, Fairbanks M. The effects of vocabulary instruction: A model-based meta-analysis. Review of Educational Research. 1986; 56:72–110.10.3102/00346543056001072

Stanovich KE. Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. Reading Research Quarterly. 1986; 21:360–406.

Stanovich, KE. Speculations on the causes and consequence of individual differences in early reading acquisition. In: Gough, P.; Ehri, L.; Treiman, R., editors. Reading acquisition. Mahwah, NJ: Erlbaum; 1992. p. 307-342.

Sternberg, RJ. Most vocabulary is learned from context. In: McKeown, M.; Curtis, ME., editors. The nature of vocabulary acquisition. Hillsdale, NJ: Erlbaum; 1987. p. 89-105.

Tannenbaum KR, Torgesen JK, Wagner RK. Relationships between word knowledge and reading comprehension in third-grade children. Scientific Studies of Reading. 2006; 10:381–398.10.1207/s1532799xssr1004_3

Thorndike, RL.; Hagen, EP.; Sattler, JM. Stanford-Binet Intelligence Scale. 4. Itasca, IL: Riverside Publishing; 1986. Technical Manual.

Torgesen JK, Wagner RK, Rashotte CA, Burgess S, Hecht S. 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. 1997; 1:161–195.10.1207/s1532799xssr0102_4

Verhoeven L, van Leeuwe J. Prediction of the development of reading comprehension: A longitudinal study. Applied Cognitive Psychology. 2008; 22:407–423.10.1002/acp.1414

Wagner RK, Meros D. Vocabulary and reading comprehension: direct, indirect, and reciprocal influences. Focus on Exceptional Children. 2010; 43:1–10.

Wagner RK, Torgesen JK, Rashotte CA, Hecht S, Barker T, Burgess S, Garon T. Causal relations between the development of phonological processing and reading: A five-year longitudinal study. Developmental Psychology. 1997; 33:468–479. [PubMed: 9149925]

Wagner RK, Torgesen JK. The nature of phonological processing and its causal role in the acquisition of reading skills. Psychological Bulletin. 1987; 101:192–212.10.1037/0033-2909.101.2.192

Wagner RK, Torgesen JK, Laughon P, Simmons K, Rashotte CA. Development of young readers' phonological processing abilities. Journal of Educational Psychology. 1993; 85:1–20.10.1037/0022-0663.85.1.83

Wagner RK, Torgesen JK, Rashotte CA. The development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. Developmental Psychology. 1994; 30:73–87.10.1037/0012-1649.30.1.73

Woodcock, RW. Woodcock Reading Mastery Tests – Revised/Normative Update. Circle Pines, MN: American Guidance Service, Inc; 1987.

Woodcock, RW.; Johnson, MB. Woodcock-Johnson Psycho-Educational Battery—Revised. Rolling Meadows, IL: Riverside Publishing; 1989.

Woodcock, RW.; McGrew, KS.; Mather, N. Woodcock-Johnson III Normative Update (NU) Complete. Rolling Meadows, IL: Riverside Publishing; 2001, 2007.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

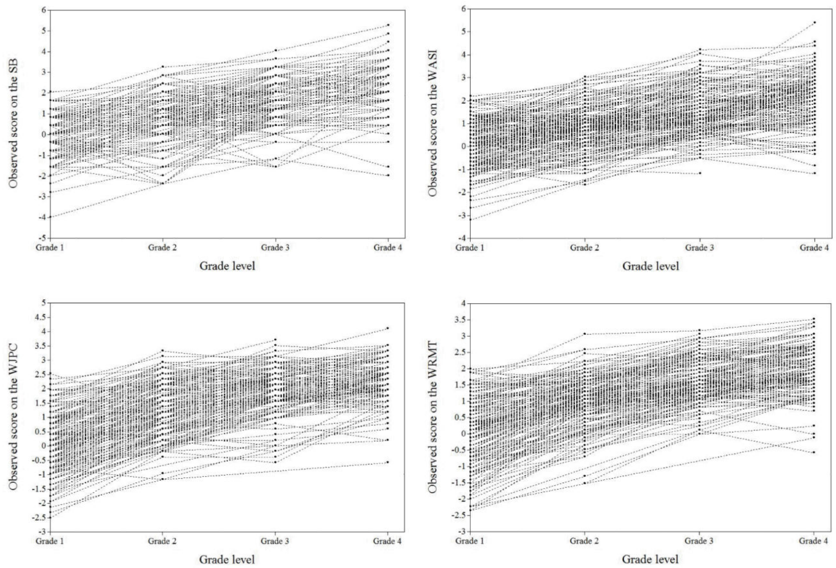

Figure 1.

Observed trajectories for the Stanford-Binet (SB) vocabulary, Wechsler Abbreviated Scales of Intelligence (WASI) vocabulary, Woodcock-Johnson Passage Comprehension (WJPC), and Woodcock Reading Mastery Test (WRMT). All values were converted to *z* scores based on means and standard deviations from Time 1.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

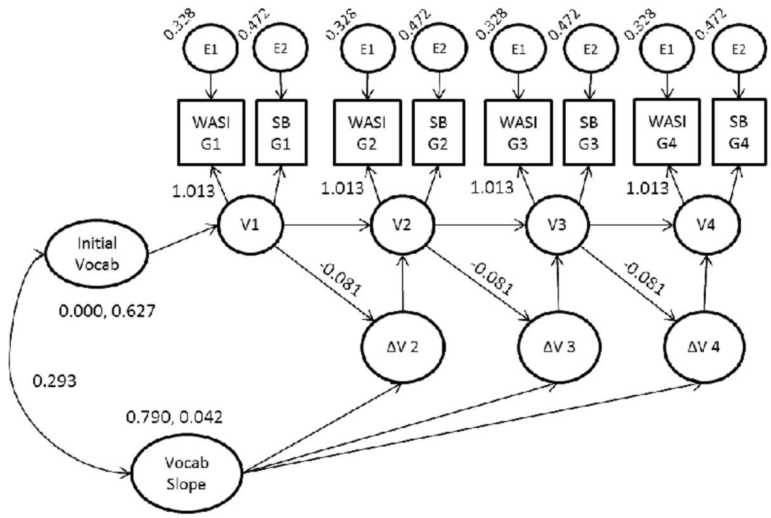

Figure 2.

Dual change score model for vocabulary knowledge. Diagram with path coefficients for the dual change score model of V. Paths with no coefficient are fixed to 1.

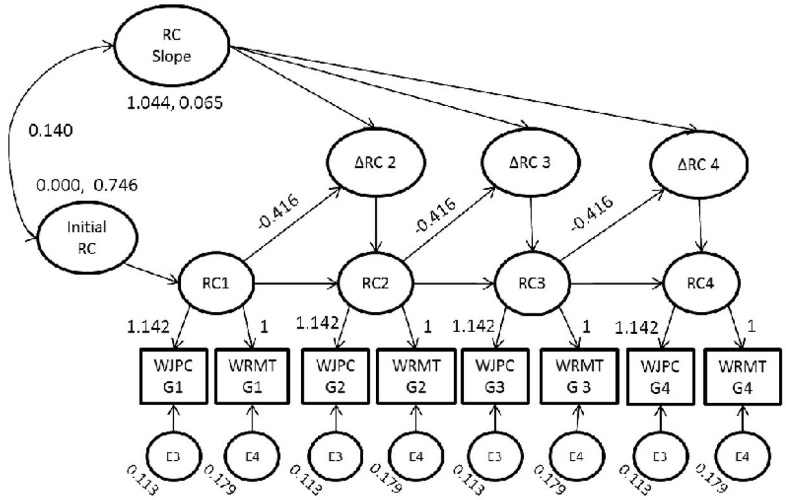

Figure 3.

Dual change score model for reading comprehension. Diagram with path coefficients for the dual change score model of RC. Paths with no coefficient are fixed to 1.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

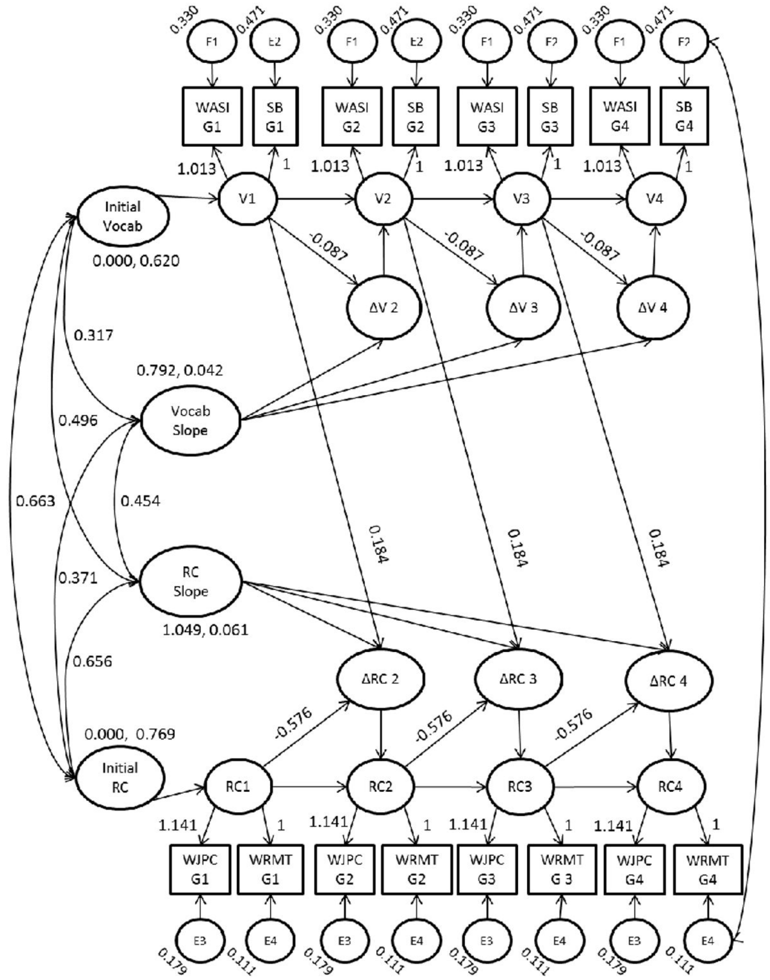

Figure 4.

Bivariate latent change score model path diagram. Diagram with path coefficients for the bivariate latent change score model with vocabulary to changes in reading comprehension coupling only. Paths with no coefficient are fixed to 1.

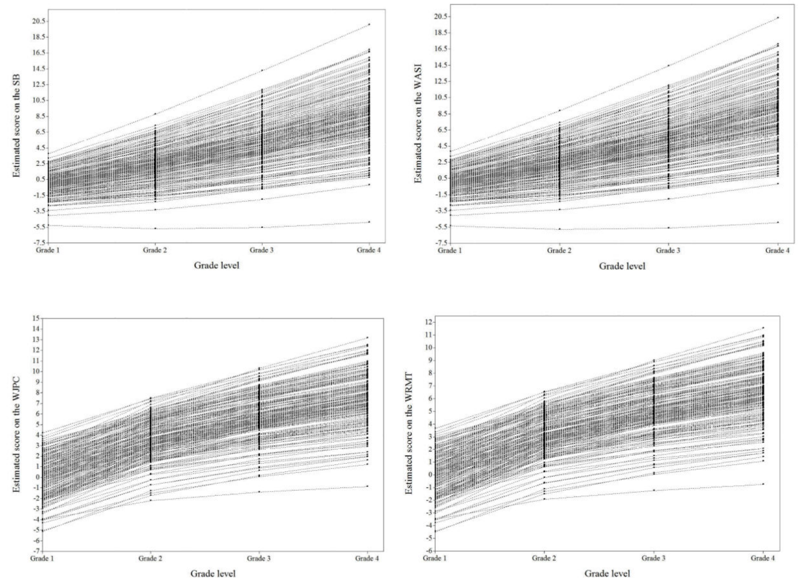

Figure 5.

Estimated trajectories for the Stanford-Binet (SB) vocabulary, Wechsler Abbreviated Scales of Intelligence (WASI) vocabulary, Woodcock-Johnson Passage Comprehension (WJPC), and Woodcock Reading Mastery Test (WRMT). All values are in the *z*-score scale.

Table 1

Sample statistics and correlations between vocabulary knowledge and reading comprehension measures

Measure	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.
1. WJPC - G1																
2. WJPC - G2	.758															
3. WJPC - G3	.698	.793														
4. WJPC - G4	.554	.639	.627													
5. WRMT -G1	.872	.763	.747	.565												
6. WRMT - G2	.735	.810	.777	.640	.758											
7. WRMT - G3	.702	.759	.791	.677	.734	.781										
8. WRMT - G4	.651	.760	.750	.715	.663	.780	.783									
9. SB - G1	.484	.504	.558	.459	.509	.518	.509	.505								
10. SB - G2	.485	.575	.582	.485	.461	.556	.578	.562	.546							
11. SB - G3	.494	.543	.584	.528	.487	.534	.533	.520	.474	.566						
12. SB - G4	.501	.523	.560	.575	.480	.519	.555	.617	.513	.569	.660					
13. WASI - G1	.512	.568	.576	.529	.526	.555	.559	.565	.675	.601	.570	.612				
14. WASI - G2	.524	.610	.666	.472	.514	.608	.646	.605	.545	.6.8	.610	.626	.656			
15. WASI - G3	.560	.608	.616	.615	.556	.580	.647	.680	.509	.583	.674	.655	.629	.732		
16. WASI - G4	.542	.582	.624	.566	.528	.589	.636	.669	.488	.624	.612	.710	.610	.675	.714	
N	315	271	260	219	316	270	260	219	315	266	260	219	314	268	260	219
Raw Scores																
Mean	18.0	24.2	27.8	29.8	21.0	29.6	35.1	38.1	18.9	20.9	22.9	24.3	24.0	28.3	33.1	37.0
SD	5.1	4.7	4.2	4.2	8.5	6.2	6.0	6.2	2.5	2.7	2.7	2.7	5.9	6.0	6.3	6.8
Minimum	5	12	14	15	1	8	16	16	9	13	15	14	5	7	17	17
Maximum	31	35	37	55	38	47	51	52	25	27	30	32	37	44	55	58
Standardized																
Mean	104.3	103.2	102	101.7	108.6	107.6	106.1	105.3	49.8	50.7	51.4	51.4	49.4	51.3	52.7	52.6
SD	13.0	10.5	9.8	9.3	10.3	9.3	9.8	10.1	6.3	6.8	6.5	5.8	9.7	9.3	9.4	9.3
Minimum	5	12	14	15	1	8	16	16	9	13	15	14	5	7	17	17
Maximum	31	35	37	55	38	47	51	52	25	27	30	32	37	44	55	58

Note. WJPC = Woodcock-Johnson Passage Comprehension, WRMT = Woodcock Reading Mastery Test, SB = Stanford-Binet Vocabulary, WASI = Wechsler Abbreviated Scales of Intelligence Vocabulary, G1 = Grade 1; G2 = Grade 2; G3 = Grade 3; G4 = Grade 4; *N* = Number of participants; SD = standard deviation; All correlations significant at $p < .05$.