Published in final edited form as:

J Educ Psychol. 2014 May 1; 106(2): 419–434. doi:10.1037/a0035692.

Developmental Relations between Reading and Writing at the Word, Sentence and Text Levels: A Latent Change Score Analysis

Yusra Ahmed,

University of Houston Texas Institute for Measurement, Evaluation and Statistics

Richard K. Wagner, and

Florida State University Florida Center for Reading Research

Danielle Lopez

Florida State University Florida Center for Reading Research

Abstract

Relations between reading and writing have been studied extensively but the less is known about the developmental nature of their interrelations. This study applied latent change score modeling to investigate longitudinal relations between reading and writing skills at the word, sentence and text levels. Latent change score models were used to compare unidirectional pathways (reading-to-writing and writing-to-reading) and bidirectional pathways in a test of nested models. Participants included 316 boys and girls who were assessed annually in grades 1 through 4. Measures of reading included pseudo-word decoding, sentence reading efficiency, oral reading fluency and passage comprehension. Measures of writing included spelling, a sentence combining task and writing prompts. Findings suggest that a reading-to-writing model better described the data for the word and text levels of language, but a bidirectional model best fit the data at the sentence level.

Keywords

reading; writing; spelling; literacy; language; latent change score

Although historically most research and pedagogy has separated reading and writing instruction (Shanahan, 2006), relations between reading and writing (i.e. literacy skills) have been studied extensively over the past couple of decades. Most studies find that reading and writing are highly related (e.g., Jenkins et al., 2004; Berninger et al., 2002; Abbott & Berninger, 1993; Tierney & Shanahan, 1996; Juel et al., 1986; Juel, 1988, 1983; Loban, 1963; Shanahan, 1984), and neuroimaging studies have shown that reading and writing activate overlapping brain regions (Pugh et al., 2006). Furthermore, interventions that have focused on transfer of skills show that reading instruction has a positive effect on writing (Shanahan, 2006) and writing instruction on reading (Weiser & Mathes, 2011; Graham &

Hebert, 2011; Tierney & Shanahan, 1991). However, less is known about the developmental nature of the interrelations.

Whereas some earlier studies examined correlations (e.g., Shanahan, 1984), other studies included multivariate structural equation modeling (SEM) and were cross-sectional in nature (e.g., Shanahan & Lomax, 1986, 1988). Structural equation models include various measures to form reading and writing latent factors. In an exploratory multivariate analysis, Shanahan (1984) examined the relation among several reading measures (phonics, reading comprehension, and vocabulary) and writing measures (vocabulary diversity, syntactic complexity, qualitative and quantitative measures of spelling and story organization) of approximately 250 students in grades 2 and 5, and showed that correlations between the reading and writing measures ranged from .14 (vocabulary and number of writing episodes) to .68 (phonics and orthographic accuracy) in grade 2, and from .13 (reading comprehension and average t-unit length) to .62 (vocabulary and orthographic accuracy as well as vocabulary and spelling) in grade 5. Shanahan and Lomax's (1986, 1988) studies are one of the earliest examples of structural equation models used to examine potential bidirectional reading-writing relations. Their studies included three factors for reading (phonetic word analysis, vocabulary, and comprehension) and four for writing (spelling, vocabulary diversity, syntax, and story structure), where reading comprehension, spelling and story structure were latent variables comprised of two or more observed variables. Their results showed that the bidirectional model provided the best fit to the data in grades two and five in comparison to the alternative unidirectional models, where reading factors influenced writing factors and writing factors influenced reading factors. The amount of shared variability between reading and writing rarely exceeded 50% in correlational studies using observed indicator variables (Shanahan, 2006). In contrast, with the use of latent factors in more recent studies, estimates of shared variability between reading and writing have ranged from 77-85% at the word level and about 65% for text level variables (Berninger et al., 2002). A recent study of the development of spelling and word-level reading in Dutch 1st through 6th graders reported divergence in their developmental patterns. Phonological awareness and knowledge of letter-sound correspondences predicted early word-level reading more than later word-level reading, but they were consistent predictors of spelling across the developmental range examined (Vaessen & Blomert, 2013)

Various approaches have been used to investigate longitudinal relations between developing reading and writing (e.g., Abbott et al., 2010; Metha et al., 2005; Lerkkanen et al., 2004). One frequently used approach is cross-lagged correlational or structural equation modeling. Cross-lagged models use longitudinal data, and examine how one variable in the model influences itself over time (i.e., auto-regressive parameters) and how each variable crosses over to influence the other variable at subsequent times (i.e., cross-regressions). Lerkkanen, Rasku-Puttonen, Aunola & Nurmi (2004) conducted a study that examined bidirectional relations between reading and writing using a sample of first grade Finnish students. Finnish is a transparent orthography and has more direct letter sound correspondences than English (i.e., students make fewer spelling errors). Their analyses included one latent variable for initial reading skills (letter naming and word list reading), one latent variable for reading (word reading and reading comprehension) and two latent variables for writing, spelling and writing fluency (measured as writing as many words or sentences, or a story about a given

picture). The model was a cross-lag SEM including bidirectional relations between the reading factor and a writing factor. Initial reading skill was added as a covariate to predict Time 1 reading and writing. After removing non-significant pathways, their results showed that reading and spelling were reciprocally related during the first semester, but in subsequent semesters, reading predicted spelling, and writing fluency predicted reading.

Although these studies and several others show that relations between reading and writing appear to be bidirectional (Shanahan & Lomax, 1986; Lerkkanen et al., 2004; Abbott et al., 2010), other studies suggest that the relation is largely unidirectional. Some studies have reported that writing influences reading (e.g., Caravolas et al., 2001; Cataldo et al., 1988; Shanahan & Lomax, 1988; Berninger et al., 2002) and other studies have reported that reading influences writing (Shanahan & Lomax, 1986; Aarnoutse et al., 2005; Babagayigit & Stainthorpe, 2011; Sprenger-Charolles et al., 2003; Berninger et al., 2002). Part of the reasons for the mixed findings from previous studies is that previous studies varied in the number and type of indicators used to represent constructs. For example, Lerkkanen et al. (2004) defined reading using a composite of word reading and reading comprehension measures, and writing fluency was broadly defined as the total number of words produced by allowing students to write either words or sentences, or a story about a picture. Shanahan and Lomax (1986) on the other hand included separate latent factors for components of reading (word analysis, vocabulary, and comprehension) and components of writing (spelling, vocabulary diversity, syntax, and story structure) in a single multivariate model.

Levels of Language Approach

Recent studies have analyzed separate components of reading and writing based on a levels of language approach that differentiates the levels of the word, sentence and passage (Abbott et al., 2010; Berninger et al., 2002; Whitaker, Berninger, Johnston, & Swanson, 1994). This approach is supported by the finding that intraindividual differences exist across levels of language (word, sentence, and text) for reading (e.g., Vellutino, Tunmer, Jaccard & Chen, 2004) as well as writing (Abbott & Berninger, 1993; Whitaker, Berninger, Johnston, & Swanson, 1994), suggesting that children could be adequate at decoding but not reading comprehension, or adequate in spelling but not sentence formation. Research on linguistics, psychology and educational sciences further suggests there are common constructs underlying literacy development. These constructs include knowledge of phonological structures, knowledge of the alphabetic principle, fluency in decoding and encoding, comprehension of oral and written language, and wide reading and writing (Foorman et al., 2011). Furthermore, reading and writing depend on common knowledge of specific components of written language that can be subdivided into grapho-phonics, text attributes of syntax and text format (Fitzgerald & Shanahan, 2000).

Decoding and encoding words

Alphabetic writing systems rely on a relatively small number of orthographic units or letters that map roughly onto the phonemes of speech. For example, the letter 's' is used to represent what actually are different speech sounds or phones associated with the 's' in 'same,' 'sure,' and 'spot.' The alphabetic principle holds that there is a rough correspondence between phonemes and the letters in an alphabetic system of writing, and

children rely on this grapheme-phoneme correspondence in order to read and write words. Fitzgerald & Shanahan (2000) suggested that writers rely on grapho-phonics, which requires phonological awareness, grapheme awareness, and morphology (Shanahan, 2006). For example, for decoding the beginning of the word “sure” a reader chooses between the potential phonological representations /s/, /z/, /sh/ or a silent letter. For encoding the same word, however, a writer chooses from the s, sh, or ch orthographic paths (Shanahan, 2006; Sprenger-Charolles et al., 2003). Most researchers suggest that encoding is not a reversal of decoding, although both rely on knowledge of the alphabetic principle (Abbott et al., 2010; Foorman et al., 2011; Shanahan, 2006).

Sentence reading and writing

The grammatical rules and punctuations used in creating sentences are attributes of syntax (Shanahan, 2006). Both readers and writers rely on meaningful syntactic orderings of words as well as the knowledge of punctuation marks to create sentence boundaries. Several studies have shown that children are sensitive to linguistic constraints in oral language as well as written language (e.g., Bates & Goodman, 1999; Rode, 1976), even at the preschool level (Puranik & Lonigan, 2010). Most of the syntactic structures found in written language are learned through many years of schooling (Beers & Nagy, 2011), during which children are also learning to read. Research on combining sentences suggests that writers first acquire syntax and semantics at the level of the phrase, but they are unable to form larger units of meaning without error (Rose, 1976). Research on syntactic complexity of writing has shown that writers use complex syntactic structures (e.g., clauses, and complex phrases), although how this development occurs is still not known as most research has focused on the development of writing at the text level or has used cumulative measures of writing (Beers & Nagy, 2011; Berninger et al., 2010). Although both reading and writing of sentences begin with developing clauses within sentences, little research has been conducted to examine the development of reading and writing at the sentence level.

Text reading and writing

Recent studies have shown that the correlations between passage comprehension and text composition range from moderate to high for both children and adults (Berninger et al., 2002; Kim, Park, & Park, under review) and that reading comprehension and composition are mutually predictive over time (Abbott et al., 2010). Readers apply a series of inferences and construct propositions based on the information provided by the text (Foorman et al., 2011; Kintsch & Mangalath, 2011). Additionally, they form mental models of the text that represent the situation described in the text. Analogous to reading comprehension, composition is also a complex process, entailing translation of ideas into writing as well as transcription skills (handwriting and spelling; Hayes & Berninger, 2010). However, the pattern of reasoning is different for each process: while readers focus on gaining support for their interpretations, writers focus on strategies to create meaning (Langer, 1986).

The important conclusion from the reading-writing research is that although reading and writing are not inverse processes, they rely on similar cognitive mechanisms that allow for simultaneous growth as well as transfer of knowledge. A recent study by Abbott, Berninger and Fayol (2010) modeled the longitudinal development of reading and writing across levels

of language. Their study placed emphasis on integrating levels of language by specifying several bidirectional models. Their study included data of children who were tested longitudinally from grades one through seven. The first bivariate model was a sub-word/word level model that included handwriting, spelling and word reading. They included handwriting (a sub-word skill) to clarify conflicting results of earlier longitudinal studies that show both bidirectional and unidirectional relations between word reading and spelling. The results showed a significant bidirectional relation between word reading and spelling across grade 2 through 7. For grade 1, however, only the spelling to reading pathway was significant.

Their second bivariate model included pathways between word (word reading and spelling) as well as text level measures (reading comprehension and written composition). The word/text model was based on the simple view of reading that holds that reading comprehension is a product of word level reading and listening comprehension. Similarly, this theoretical framework was applied to writing, whereby writing was conceptualized as a product of word level writing (spelling) and text level writing (composition). Similar to the first model, the results for this model showed significant bidirectional relations between word spelling and reading for grade 2 through 7. For grade 1, however, only the path from word reading to spelling was significant in this model. Furthermore, for both sub-word/word and word/text level models the magnitude of the univariate spelling-spelling and reading-reading stability parameters (i.e., autoregressors) was larger (range = .59 to .93) than spelling-reading and reading-spelling parameters (range = .14 to .33). Similarly, at the text level, their results showed that the magnitude of the bivariate parameters were small. Specifically, reading comprehension predicted composition in grades 2 to 6 (range = .13 to .22) and composition predicted comprehension in grades 3 to 5 (range = .18 to .20). The stability parameters were larger in magnitude and ranged from .47 to .62 for reading comprehension, and .26 to .41 for composition.

Overall, their results indicate relations between reading and writing at the word and text levels are less clear in grade 1 than in subsequent grades, possibly because other factors (such as verbal ability or exposure to print) contribute to the development of reading and writing at that grade level. The results showed that differences in magnitude between the stability parameters and bivariate parameters show a weak relation between reading and writing processes, and strong a reading-reading and writing-writing relation. However, their study used observed indicators as measures of reading and writing ability. The use of observed indicators assumes that the constructs were measured without error, and the stability parameters could largely be inflated due to common method variance.

Latent Change Score Modeling

A new model for analyzing longitudinal data called latent change models provide an important new tool for analyzing relations between developing reading and writing (McArdle, 2009). Latent change score (LCS) models have been applied to various areas of research, including human development, social and personality psychology, psychopathology, cognitive psychology and neuropsychology to examine dynamics among developmental processes (Ferrer & McArdle, 2007; McArdle & Grimm, 2010). LCS models

combine the strengths of cross-lagged structural models and latent growth curve models, the two kinds of models that are most commonly applied to longitudinal developmental data. Similar to cross-lagged structural models and unlike latent growth curve models, LCS models divide development into segments of time. This allows the use of time-sequence logic (e.g., does early development of reading account for later development of writing) to test alternative models of developmental influence among constructs. Similar to latent growth curve models and unlike cross-lagged structural models, LCS models explicitly model growth in absolute performance over time. Thus, all of the available data, means as well as covariances, are used to test alternative models¹.

The LCS models used in the present study describe group average change over time. Unlike growth curve models, the growth estimates are not based on slopes (or rate of change) but on differences between two scores. The changes () are essentially the subtraction between the common factor at a time point (t) and the previous time point ($t - 1$) (McArdle, 2009). Estimating latent change scores allows us to examine whether current level of either reading or writing explains later change in reading or writing. It also is possible to examine whether early change in either reading or writing explains later change (Grimm et al., 2012). Univariate parameters (e.g., reading-reading) are pathways a and b in Figure 1, and multivariate parameters (e.g., reading-writing) are pathways c and d . In the interest of parsimony, the pathways were constrained to be the same over time. Thus, the same dynamics are at play during all time periods (Jajodia, 2012). Additionally, changes are accumulated over time, such that the last change score is an addition of all the previous status and change scores. For example, if we measured a child's weight instead of an academic skill, the last change score would be an estimate of all the weight gained and lost during the span of four years.

Method

Participants and Instructional Context

The first year sample consisted of 316 first-grade children (49% female), and ranged in age from 6 years, 1 month to 8 years, 9 months ($M = 6.6$, $SD = 0.56$). The sample was representative of the student population in Florida, with 60% Caucasian, 25% African American, 4% Hispanic, 4% Asian and 7% another ethnicity.

In the second year of data collection, 270 children (48.5% female) were still included in the study. The second year sample ranged in age from 7 years, 2 months to 9 years, 5 months ($M = 7.5$, $SD = .67$). The ethnic composition of the sample was 63% Caucasian, 21.5% African American, 5.2% Hispanic, 4.4% Asian, 5.2% mixed ethnicity and 0.7% another ethnicity. A small number of students ($n=12$) were retained in grade 1 during year 2 of the study.

In the third year of data collection, 260 children (48.8% female) remained in the study. The third year sample ranged in age from 8 years, 11 months to 10 years, 11 months ($M = 8.5$,

¹Z-scores are different from latent change scores because z-scores centralize the data relative to a sample mean (in this study we used the mean of Grade 1), and distributions expressed in z-scores have a mean of 0 and standard deviation of 1 (see Grade 1 in Table 1). Latent change scores represent individual differences in true scores between a time point and the previous time point, and each latent change score has its own distribution (see Table 5).

$SD = .56$). The ethnic composition was 61.5% Caucasian, 23.5% African American, 5% Hispanic, 5.8% Asian and 4.2% mixed ethnicity. A small number of students ($n=2$) were retained in grade 2 during year 3 of the study.

In the fourth and final year of data collection, 219 children (48.1% female) remained in the study. The fourth year sample ranged in age from 9 years, 10 months to 11 years, 1 month ($M = 9.9$, $SD = .42$). The ethnic composition was 59.1% Caucasian, 24.2% African American, 4.7% Hispanic, 4.7% Asian, 5% mixed ethnicity and 1.2% unknown. A small number of students ($n=13$) were retained in grade 3 during year 4 of the study.

Following the Sunshine State Standards, students are required to take the Florida Comprehensive Assessment Test (FCAT) yearly. The FCAT contains a Reading subtest which assesses informational and literary reading comprehension. The Writing subtest requires students to write a narrative, expository or persuasive essay within a 45-minute session. Accordingly, schools in Florida provide a minimum of 90 minutes of instruction in arts and literacy, including instruction on phonological awareness, phonics, fluency, vocabulary, and comprehension, although many classrooms/ schools exceed 90 minutes. Furthermore, schools in the Leon County District, where this study was conducted, use a reading program called Imagine It! (formerly Open Court). This reading program takes a balanced approach to instruction that combines explicit instruction in skills like phonological awareness and phonics with considerable opportunity to acquire skills implicitly through reading. Writing instruction focuses on improving key components of writing, including focus, organization, support, and conventions (FLDOE, 2013).

Measures

Decoding—Two forms of the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999) Phonetic Decoding Efficiency (PDE) subtest were used to assess pseudo-word reading fluency and accuracy. Forms C and D were used during years 1 and 3 and Forms A and B were used during years 2 and 4 to form a pseudo-word decoding latent factor. The TOWRE PDE requires accurately reading as many non-words as possible in 45 seconds. The TOWRE test manual reports test-retest reliability to be .90 for the PDE subtest.

Sentence Reading—Two forms of the Test of Silent Reading Efficiency and Comprehension (TOSREC, forms A and O; Wagner et al., 2010) were used as measures of silent reading fluency. Students were required to read brief statements (e.g., “a cow is an animal”) and verify the truthfulness of the statement by circling yes or no. Students are given three minutes to read and answer as many sentences as possible. The mean alternate forms reliability for the TOSREC ranges from .86 to .95. The TOSREC is designed to measure both silent reading fluency and comprehension.

Oral Reading Fluency—Two forms of the *Dynamic Indicators of Basic Early Literacy Skills (DIBELS)* Oral Reading Fluency (Good & Kaminski, 2001) were used as indicators of text reading fluency. Students were asked to read a passage out loud for 1 minute. The number of correct words from the passage is the total score. Words omitted, substituted, and hesitations of more than three seconds were scored as errors. Test-retest reliabilities for elementary students are reported to range from .92 to .97.

Reading Comprehension—The *Woodcock-Johnson-III* (WJ-III) Passage

Comprehension subtest (Woodcock et al., 2001) and the *Woodcock Reading Mastery Test-Revised* Passage Comprehension subtest (WRMT-R; Woodcock, 1987) were used to provide two indicators of reading comprehension. For both of the passage comprehension subtests, students read brief passages to identify missing words. Testing is discontinued when the ceiling is reached (six consecutive wrong answers or until the last page was reached). According to the test manuals, test-retest reliability is reported to be above .90 for WRMT-R, and the median reliability coefficient for WJ-III is reported to be .92. Furthermore, it's possible that reading comprehension tests measure different skills depending on developmental level, reflecting a greater influence of decoding for younger and less skilled children than older and more skilled children (see Keenan, Betjemann, & Olson, 2008).

Spelling—The Spelling subtest from the *Wide Range Achievement Test-3* (WRAT-3; Wilkinson, 1993) and the Spelling subtest from the *Wechsler Individual Achievement Test-II* (WIAT-II; The Psychological Corporation, 2002) were used to form a spelling factor. Both spelling subtests required students to spell words with increasing difficulty from dictation. The ceiling for the WRAT3 Spelling subtest is misspelling ten consecutive words. If the first five words are not spelled correctly, the student is required to write his or her name and a series of letters and then continue spelling until they have missed ten consecutive items. The ceiling for WIAT-II is misspelling 6 consecutive words. The reliability of the WRAT-3 spelling subtest is reported to be .96 and the reliability of the WIAT-II Spelling subtest is reported to be .94.

Written Expression—The Written Expression subtest from the *Wechsler Individual Achievement Test-II* (WIAT-II; The Psychological Corporation, 2002) was administered. Written Expression score is based on a composite of Word Fluency and Combining Sentences in first and second grades and a composite of Word Fluency, Combining Sentences, and Paragraph tasks in third grade. In this study the Combining Sentences task was used as an indicator of writing ability at the sentence level. For this task students are asked to combine various sentences into one meaningful sentence. According to the manual, the test-retest reliability coefficient for the Written Expression subtest is .86.

Writing Prompts—*Writes Upon Request* (WUR) is a district level program that assesses student writing in first through tenth grades, and is used to prepare students for the Writing portion of the FCAT. Trained classroom teachers administered the tests and scored the samples based on a holistic score of 1 to 6 (FLDOE, 2013). The holistic score was based on a) focus, or the ability to maintain an idea, theme or unifying point, b) organization, or the structure and relationship between the beginning, middle and end of the text, c) support, or the quality of details used to explain, clarify and define and d) conventions, or punctuation, capitalization, spelling and sentence structure. Information about the reliability of scoring was not available. Because the writing quality score used in this study was an assessment of overall writing quality, it's possible that different results might hold when specific components of written expression are scored separately, for example, macro-organization, voice, how well arguments are made, etc. A second writing composition task was also administered. Participants were asked to write a passage on a topic provided by the tester for

10 minutes. Students were instructed to scratch out any mistakes and were not allowed to use erasers. Both the WUR and compositional fluency prompts included narrative and expository prompts. The topics for the compositional fluency task were as follows: choosing a pet for the classroom (grade 1), favorite subject (grade 2), a day off from school (grade 3), a time you went on a field trip (grade 4). A computational index of total number of words produced was used as an indicator for text level writing. Reliability is nearly perfect for the total number of words as it is an automated score and requires no human judgment. Nelson and Van Meter (2007) report that total word productivity is a robust measure of developmental growth in writing.

Procedures

Participants were tested on all measures once a year, approximately one year apart. Participants were first grade students in the fall of 2007 whose parents consented to participate in the longitudinal study. Participants attended six different schools in a metropolitan school district in Tallahassee, Florida. Data were gathered by trained testers during thirty to sixty minute sessions in a quiet room designated for testing at the schools.

Results

Prior to analysis, data were screened for missing values, outliers, and normality. The attrition rate from the previous year was 14.5% ($n = 46$) at Time 2, 3.7% ($n = 10$) at Time 3 and 15.8% ($n = 41$) at Time 4. Reasons for not being included in years 2-4 were moving out of the area, no longer wishing to participate, or unable to be contacted. Students who dropped prematurely from the study did not differ significantly from those who completed the study on demographics and on most variables at Time 1. Significant differences were found for the DIBELS Oral Reading Fluency tests, $F[1,314] = 4.07$, $p = 0.04$ and $F[1,314] = 4.09$, $p = 0.04$ respectively. These small differences due to attrition bias present threats to internal validity and should be considered when interpreting the results for the text-level fluency model. Full-information maximum-likelihood estimation (FIML) was used in Mplus 7 (Muthén & Muthén, 1998-2013) to handle missing data points. This approach was utilized because maximum likelihood estimates of missing data provide the least biased estimates (Little & Rubin, 1989).

Outliers were corrected using the median $\pm$ two interquartile range as the criterion for classifying a data point as an outlier (Tabachnik & Fidell, 2007). As a final step, each child's scores were standardized based on their year 1 scores. The z score standardizations indicate how many standard deviations a child's score is above or below their year 1 score¹. Z scores are useful for comparing values of variables that are measured on different scales. Table 1 contains descriptive statistics for all measures. Evaluation of skewness and kurtosis statistics revealed mild departure from normality for Grade 2 WRMT Passage Comprehension (kurtosis = 2.49) and Grade 3 WRMT Passage Comprehension (kurtosis = 2.92), likely due to few data points hitting the floor on these variables ($n=3$ for grade 2; $n=1$ for grade 3). The remaining skewness and kurtosis values fell within an acceptable range indicating that the data were normally distributed. Tables 2 and 3 present correlations reported by year.

Test of Nested Models

A series of latent change score models were conducted using M-plus 7 (Muthen & Muthen, 1998-2013). Three nested models (two unidirectional and a bidirectional model) were tested for each level of language (word, sentence and text). In the reading-to-writing model (see Figure 1, 3 and 4), the pathways from reading to writing were estimated whereas the pathways from writing to reading were fixed at zero. In the writing-to-reading model the pathways from writing to reading were estimated and the pathways from reading to writing were fixed at zero. In the bidirectional model (see Figure 2) all pathways were estimated. The stability parameters (auto-regressions between true scores) and the pathway from the latent change score to its true score were fixed at 1 to meet model identification requirements (McArdle, 2009; McArdle & Grimm, 2010). Univariate pathways (called auto-proportions; pathway *a* in Figure 1) are regression coefficients of the change score regressed on its true score at the previous time point. Univariate parameters also include auto-regressions between changes² (pathway *b* in Figure 1). Multivariate pathways (called coupling effects, or status to change) include the change scores regressed on the other true score at an earlier time point (pathway *c* in Figure 1) as well as the change score regressed on the other change score (pathway *d* in Figure 1). These additional pathways were added to estimate causal effects of change scores on subsequent change scores (see Grimm et al., 2012). Specifically, univariate auto-regressors and multivariate cross-regressions were added between the change scores. The additional impact of these pathways allowed the change in one variable to impact the change in the other. For latent change score models in this study, initial change scores are different from subsequent change scores because of the additional parameters that predict the second and third change scores. Thus, all change scores are predicted by the reading and/or writing latent factors at the previous year, but the second and third change scores are also predicted by previous change scores. The correlation between the latent factors at Year 1 was included to account for the action of causal effects prior to the time period modeled. The double-headed arrows going from a variable to itself represent the variances of Year 1 true scores, as well as residual variances of errors and the residual variance of the change scores. Strict measurement invariance was established by setting the loadings of the factor to be equal over time, and the residual variances of the observed variables were also set to be equal over time. This was done in order to ensure the true scores measured the same latent skill over time.

As the unidirectional models are nested within the bidirectional model, chi-square difference testing was used to statistically compare models. The base model, or the bidirectional model, always yields the best fit as it has the fewest constraints and is thus the least restrictive. However, in the interest of parsimony the most restrictive nested model that is not significantly worse fitting is preferred. Fit indices for the constrained models are presented in Table 4.

Comparison of the unidirectional and bidirectional models with a χ^2 difference test revealed that the word and text level reading-to-writing models did not provide a significantly poorer fit to the data than did the less-constrained bidirectional model. Thus, they were considered

²In LCS models auto-regressions between true scores are not freely estimated.

the most appropriate models at the word and text levels of language. At the sentence level, a χ^2 difference test revealed that the unidirectional model provided a significantly poorer fit than did the bidirectional model. Consequently, the bidirectional model was considered the most appropriate model only for the sentence level.

The estimated means, standard deviations and minimum and maximum values of the change scores are summarized in Table 5. Overall, the mean of the changes were positive and indicated the largest gains in all reading variables as well as spelling were made between grades 1 and 2. The largest gains in sentence writing and compositional fluency were made between grades 2 and 3. The variances of the change scores indicated there was small amount of interindividual variability in changes in reading as well as spelling, but the large variances of the sentence and text level changes indicate that student's growth trajectories varied more at these higher levels of written language.

Evaluations of Structural Models

Word Level—Figure 1 contains the unstandardized regression coefficients for the word level reading-to-writing model. Factor correlations at year 1 showed that decoding was highly correlated with spelling ($r = .72, p < 0.001$). Decoding changes were predicted by previous decoding changes ($\beta = .14, p < 0.05$), but were not predicted by decoding status ($\beta = -.03, p > .05$). The positive and significant effect of the auto-regressions indicates that children who improved on decoding continued improving on decoding across years. The non-significant value of the auto-proportions –could potentially be due to low ceiling effects that are characteristic of decoding, rather than the lack of a true relation.

Spelling changes were predicted by spelling status ($\beta = -.09, p < 0.05$), but were not predicted by spelling changes ($\beta = .04, p > 0.05$). The negative effect of the auto-proportion indicates that children who started out with lower scores on spelling improved more.

Turning to bivariate effects of reading on writing, spelling changes were predicted by decoding status at the previous year ($\beta = .13, p < 0.05$), indicating that children who scored high on decoding grew faster on spelling between years. Spelling change was also predicted by decoding change ($\beta = .13, p < .001$). The positive change-to-change effect indicates that decoding is a leading indicator of spelling, as children who grow on non-word reading can be expected to also grow on spelling. The coupling change-to-change effect is the most important effect in terms of establishing reading-to-writing relations, as it shows that gains in one variable can be used to influence subsequent gain in another.

Sentence Level—Figure 2 contains the unstandardized regression coefficients for the sentence level bidirectional model. The correlation of .30 ($p < .001$) indicates a moderate relation between the initial reading and writing factors. Sentence reading changes were predicted by sentence reading at the previous year ($\beta = -.19, p < .001$). Specifically, children who scored low on sentence reading subsequently improved on sentence reading. Similarly, sentence writing changes were predicted by sentence writing status ($\beta = -.48, p < .001$).

In terms of bivariate effects, sentence writing changes were predicted by sentence reading status ($\beta = .79, p < .001$). These findings suggest that children who scored high on sentence

reading improved on sentence writing across years. Furthermore, sentence writing status also had a small but significant effect on sentence reading changes ($\beta = .05, p < .001$), indicating that children who scored high on sentence writing improved on sentence reading across years.

Text Level—Figure 3 and 4 contains the unstandardized regression coefficients for the fluency and discourse text level models, respectively. The correlation of $.37 (p < .001)$ indicated a moderate relation between the initial reading and writing fluency, as did the correlation between reading comprehension and writing quality ($r = .49, p < 0.001$). The pattern of results for the auto-proportions was similar to the sentence level model. Results showed that reading changes were negatively predicted by reading status for fluency ($\beta = -.08, p < .001$) as well as discourse ($\beta = -.19, p < .001$). Similarly, writing changes were a function of writing status for the fluency model ($\beta = -.23, p < .001$), and the discourse model ($\beta = -.63, p < .001$). Auto-regressions were significant between writing changes in the fluency model ($\beta = -.27, p < .001$) and between reading changes in the discourse model ($\beta = -.35, p < .001$). The negative effect of the auto-regressions shows that children who made larger gains were those who had made less gains between the previous school years.

In terms of the bivariate effects, changes in writing were predicted by reading status for the fluency model ($\beta = .19, p < .05$) and the discourse model ($\beta = .48, p < .001$), suggesting that students who are adequate at text-level reading are able to make gains in writing quantity as well as quality.

In sum, univariate relations (reading-to-reading and writing-to-writing) were characterized by low status at the previous year predicting subsequent changes in reading and writing at all levels of language (with the exception of word level reading). Hence, as expected, children who were growing were those who had a low status on reading or writing achievement the previous year. Our findings also indicated that writing changes were a function of reading status, and this relationship was stronger for sentence writing and writing quality than for spelling and compositional fluency. In addition, improvement in decoding predicted an improvement in spelling across years. Finally, status in sentence writing had a small effect on the improvement in sentence reading.

Gender Differences

We further compared parameters estimates across gender because disparities in developmental patterns of writing are often found in the literature. The means of the change scores were similar for both genders across constructs, except for writing quality (see Table 6). In terms of univariate and multivariate parameters, our results showed differences only in the auto-regressive pathways, such that word reading ($\beta_{\text{female}} = .18, p < .05; \beta_{\text{male}} = .09, p > .05$) and reading comprehension ($\beta_{\text{female}} = -.47, p < .05; \beta_{\text{male}} = -.016, p > .05$) auto-regressions were significant for females but not males, and sentence reading ($\beta_{\text{male}} = .17, p < .05; \beta_{\text{female}} = .05, p > .05$) and compositional fluency ($\beta_{\text{male}} = -.23, p < .05; \beta_{\text{female}} = -.12, p > .05$) were significant for males but not females. These results suggest that girls who make gains in decoding continue to make gains in decoding. On the other hand, girls who make gains in reading comprehension are those who didn't initially make gains in reading

comprehension. Boys who made gains in sentence reading continue to make gains in sentence reading, and boys who made gains in compositional fluency are those who didn't initially make gains in compositional fluency. These results suggest that reading-writing relations are similar for girls and boys, but the developmental pattern of writing quality may be somewhat different. Finally, we examined the residual variances of the change scores and found there was sizable variance left to be explained, especially for sentence writing, compositional fluency, and to a lesser extent writing quality.

Discussion

The purpose of the current study was to investigate the unidirectional and bidirectional relations between reading skills (decoding, sentence reading, text reading fluency and text comprehension) and writing skills (spelling, writing sentences, compositional fluency, writing quality). The use of latent change score modeling allowed us to investigate the nature of the improvement in reading and writing across grades 1 through 4, and examine the possible contributions of the various reading processes as leading indicators of growth in writing. The reverse relation (writing-to-reading) and bidirectional relations were also examined. The best fitting models were then examined in the context of higher-order cognitive and linguistic factors.

Although research supports the existence of bidirectional relations between reading and writing (Lerkkanen et al., 2004; Shanahan & Lomax, 1986, 1988; Abbott et al., 2010), the results of latent change score modeling used in the present study were that reading-to-writing models were superior to writing-to-reading and bidirectional models, especially for the word and text levels of writing. At the sentence level, a bidirectional model was superior although the writing-to-reading pathways were very small. Our findings suggest that reading exerts a relatively larger influence on writing factors than the influence of writing on reading factors. In other words, our results show that reading and writing are related (and more so at the word level) and that children apply the knowledge base used in reading to their writing across all levels of language, but this developmental pattern is not reversible (i.e., with the exception of sentence reading, children do not apply their knowledge of writing to improve their reading). This finding is in line with the view that reading and writing rely on a similar knowledge base, but they are neither reversible nor identical processes (Shanahan, 2006; Abbott et al., 2010; Foorman et al., 2011) as teaching one skill independently (reading or writing) is not automatically applied in the context of the other skill. In this sense, reading and writing are separate processes that have unique properties. However, our findings as well as interventions that are based on combinations of reading and writing (e.g., see Graham & Hebert, 2011) indicate that learning a counterpart skill from reading or writing could have cross-modal benefits that are based on procedural knowledge rather than shared semantic knowledge.

The correlations between reading and writing at year 1 were of particular interest. The relation was strong at the word level ($r = .72$) and moderate at the sentence level ($r = .30$), text-fluency level ($r = .37$) and text-discourse level ($r = .49$). This may reflect that the relation between reading and writing decreases as a function of complexity of language. This is consistent with the research by Berninger and colleagues that show a higher

correlation between word recognition and word spelling than for text level variables for both children and adults (Abbott & Berninger, 1993; Berninger, Vaughan et al., 1998; Berninger et al., 2002). This finding is also consistent with research by Metha and colleagues (2005) who suggest that decoding and spelling form a common literacy factor at the word level.

The findings of this study showed that it's possible to disentangle the variability in change from overall achievement status, and that both status and change play an important role as predictors. As expected, the results from the LCS analysis suggested that changes in reading and writing were characterized by a rise in scores (with the exception of writing quality for males), although the variances of the change scores suggested variability in the pattern of changes. Writing changes were predicted by achievement status and/or growth (change) in reading. At the word level, change in spelling was predicted by status in decoding, suggesting that skilled readers improve on spelling more than less skilled readers. Change in spelling was also predicted by decoding change, suggesting that children who improved on decoding between grades improved on spelling between subsequent grades. These findings are in line with the theories of reading and spelling development that emphasize the role of phonological skills in spelling development (e.g., Juel, 1988; Shanahan & Lomax, 1986; Aarnoutse et al., 2005; Babagayigit & Stainthorp, 2011; Sprenger-Charolles et al., 2003; Berninger et al., 2002), and suggest that the ability to read words correctly may facilitate writing them correctly, via mastery of phoneme-grapheme relations that are learned through reading (Ehri, 2005). The finding that an improvement in decoding leads to an improvement in spelling is also consistent with spelling interventions that are based on word and pseudo-word recognition (Shanahan, 2006).

Unlike previous research that showed that silent sentence reading fluency did not explain any unique variance in sentence combining (Berninger et al., 2011), we found that change in sentence combining was a function of high status in sentence reading at the previous year. Thus, our findings suggest that the ability to read sentences facilitates writing them. One possible explanation is that an individual who is fluent at reading sentences is more familiar with sentence structures and syntactic knowledge compared to an individual who is not fluent. This is consistent with research on combining sentences which suggests that sentence construction requires considerable cognitive effort as it is dependent on word choice, syntax, clarity, and rhythm (Saddler & Graham, 2005). The small effect of sentence writing on sentence reading suggests that writing sentences correctly also facilitates reading them because combining sentences requires knowledge of syntax and structures of sentences, which in turn facilitates reading sentences.

At the text-fluency level, reading fluency status predicted writing fluency, suggesting that children who are skilled at decoding connected text also write faster (or produce more text in ten minutes). The finding that growth in reading fluency facilitates growth in writing fluency suggests that teaching fluent decoding of connected text can lead to more writing. At the text-discourse level, status in reading comprehension predicted change in writing quality. Children who read for comprehension are more familiar with the format of larger texts and story structures, and it's possible that skilled readers apply this knowledge to their writing. .

Although some aspects of written language are likely to be specific to reading, other aspects are likely to be general to linguistic and cognitive factors. In comparison to reading, there has been considerably little research about which language and cognitive factors contribute to early writing development, but findings based on the cognitive processes of writing (e.g., planning, translating, reviewing, & revising; Hayes & Flower, 1980; Hayes, 1996) suggest key constructs identified by these theories such as oral language, verbal IQ and working memory (Shanahan, 2006). Hence, more research is required to determine the exact nature of the writing-to-reading relations in the context of language and higher order cognitive functions³. Furthermore, it's possible that other factors that are more closely related to the writing process are relevant predictors of overall writing quality. These factors include reasoning and planning, knowledge of story structure and genre, as well as grammar, syntax and punctuation rules.

Our results need to be interpreted in the context of the general education framework. Firstly, reading instruction is prioritized over writing instruction in the United States. Thus, it is possible that in the presence of rigorous writing instruction, bidirectional or writing-to-reading models may be accurate. Secondly, reading and writing abilities are normally distributed in the general education framework. It's possible that changes might be characterized by decline in scores over time if children with learning disability are examined. Thirdly, we examined reading-writing relations in English, and it is possible that the relations are different for shallower and deeper orthographies. Hence, our results apply to typically developing students from a middle class SES background. Finally, conclusions from this study have to be drawn with caution because the LCS models used in this study have been introduced only recently in the literature (McArdle, 2009; McArdle & Grimm, 2010). Nevertheless, statistical theory suggests that LCS models provide for a valid description of longitudinal data (Ferrer & McArdle, 2010).

Limitations

Finally, caution is warranted in interpreting our results due to three limitations of our study. First, we lacked specific information about the instructional context of the classrooms. It would be important in future research to use videotape or other classroom observational methodologies to better characterize the specific instructional practices that were used to teach reading and writing.

Second, our results are completely dependent on the variables included in our models. Most of the measures were brief, easily administered tasks due to the requirement of measuring performance at multiple levels of language in a battery with a practical length for repeated administration in a longitudinal study. It is important to replicate and extend these results in subsequent studies using alternative measures of reading and writing at the word, sentence,

³Preliminary analyses for adding linguistic and cognitive covariates to the LCS models presented here were also conducted. We individually added lexical access, vocabulary, working memory and listening comprehension as covariates to the best fitting models, to test whether relations between reading and writing remain significant after adding these covariates. Although an extensive presentation of these results is outside the scope of this paper, the overall pattern of results suggested that reading-writing relations found at all levels of language remained the same. Furthermore, there was considerable variance left to be explained for the sentence level of writing, suggesting that changes at higher levels of written expression might require predictors that are specific to the complexity of sentences rather than language and cognitive predictors.

and text levels. It would also be useful to include other measures that represent hypotheses about relations between reading and writing to see whether they mediate observed coupling influences on their co-development. In addition, although handwriting (a sub-word process) was indirectly measured in our study via the spelling measures (e.g., writing letters from dictation) research shows that handwriting plays an important role in the performance of written expression, even for older students (Connelley, Dockrell & Barnett, 2005). Thus, just as the explicit incorporation of handwriting measures is important for studying the development of writing, it's important to include handwriting in future studies examining reading-writing relations.

Third, we lost approximately one third of our sample due to attrition as we followed our participants from first through fourth grade. It was the case that the profiles of the children who dropped out of the study were similar to those who participated in all waves. Furthermore, the use of Full Information Maximum Likelihood (FIML) for data that is missing at random works well in these cases. Nonetheless, it is important for future studies to incorporate additional participants at each wave of data collection in order to maintain a stable sample size across waves.

In conclusion, this study extends the literature on developmental relations between reading and writing with three findings. First, we have shown that a reading-to-writing model is more adequate than a bidirectional model, especially at the word and text levels. Second, latent change score models are adequate for describing growth in reading and writing, and relations between them. Third, writing was a function of reading at all levels of language, and reading was a function of writing at the sentence level. Our results show that in order to understand the development of written expression, sentence level writing should be studied more extensively (this is evident because reading-to-writing relations were the strongest at the sentence level, and because of the large residual variances for sentence level writing). Finally, if future research corroborates that reading is a determinant of writing, interventions may benefit from exploring this relation at the word, sentence and text levels.

Acknowledgments

Support for this research was provided by Grant P50 HD052120 of the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The views stated herein are solely of the authors and do not represent the views of the NICHD.

References

- Abbott RD, Berninger VW. Structural equation modeling of relationships among developmental skills and writing skills in primary- and intermediate-grade writers. *Journal of Educational Psychology*. 1993; 85(3):85, 478–508.
- Abbott RD, Berninger VW, Fayol M. Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology*. 2010; 102(3): 281–298.
- Aarnoutse C, van Leeuwe J, Verhoeven L. Early literacy from a longitudinal perspective. *Educational Research and Evolution*. 2005; 11(3):253–275.
- Babayigit S, Stainthorp R. Modeling the relationships between cognitive-linguistic skills and literacy skills: New insights from a transparent orthography. *Journal of Educational Psychology*. 2011; 103(1):169–189.

- Bates, E.; Goodman, J. The emergence of grammar from the lexicon.. In: MacWhinney, B., editor. The emergence of language. Lawrence Erlbaum Associates; Mahwah, NJ: 1999. p. 27-80.
- Beers SF, Nagy WE. Writing development in four genres from grade three to seven: syntactic complexity and syntactic generation. *Reading and Writing*. 2011; 24:183–202.
- Berninger VW, Abbott RD, Abbott SP, Graham S, Richards T. Writing and reading: Connections between language by hand and language by eye. *Journal of Learning Disabilities*. Special Issue: The Language of Written Language. 2002; 35(1):39–56.
- Berninger V, Abbott R, Swanson HL, Lovitt D, Trivedi P, Lin S, Amtmann D. Relationship of word- and sentence-level working memory to reading and writing in second, fourth, and sixth grade. *Language, Speech, and Hearing Services in Schools*. 2010; 41:179–193.
- Berninger V, Nagy W, Beers S. Child writers' construction and reconstruction of single sentences and construction of multi-sentence texts: contributions of syntax and transcription to translation. *Reading and Writing*. 2011; 24:151–182. [PubMed: 21383865]
- Caravolas M, Hulme C, Snowling M. The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language*. 2001; 45:751–774.
- Cataldo S, Ellis N. Interactions in the development of spelling, reading and phonological skills. *Journal of Research in Reading*. 1988; 11:86–109.
- Connelly V, Dockrell JE, Barnett J. The slow handwriting of undergraduate students constraints overall performance in exam essays. *Educational Psychology*. 2005; 25(1):99–107.
- Dell GS. A spreading-activation theory of retrieval sentence production. *Psychological Review*. 1986; 93(3):283–321. [PubMed: 3749399]
- Ehri, LC. Development of sight word reading: Phases and findings.. In: Snowling, MJ.; Hulme, C., editors. *The science of reading*. Blackwell; Oxford: 2005. p. 135-155.
- Ferrer E, McArdle JJ. Longitudinal modeling of developmental changes in psychological research. *Current Directions in Psychological Science*. 2010; 19(3):149–154.
- Fitzgerald J, Shanahan T. Reading and writing relations and their development. *Educational Psychologist*. 2000; 35(1):39–50.
- Florida Department of Education [FLDOE]. 2013. Retrieved from: <http://fcad.fldoe.org/>
- Foorman BR, Arndt EJ, Crawford EC. Important constructs in literacy learning across disciplines. *Topics in Language Disorders*. 2011; 31(1):73–83.
- Good, RH.; Kaminski, RA.; Smith, S.; Laimon, D.; Dill, S. *Dynamic Indicators of Basic Early Literacy Skills*. 5th ed.. University of Oregon; Eugene, OR.: 2001.
- Graham S, Hebert M. Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*. 2011; 81(4):710–744.
- Grimm KJ, An Y, McArdle JJ, Zonderman AB, Resnick SM. Recent changes leading to subsequent changes: Extensions of multivariate latent difference score models. *Structural Equation Modeling*. 2012; 19:268–292. [PubMed: 23637519]
- Hayes, JR.; Berninger, V. Relationships between idea generation and transcription: How act of writing shapes what children write.. In: Brazerman, C.; Krut, R.; Lunsford, K.; McLeod, S.; Null, S.; Rogers, P.; Stansell, A., editors. *Traditions of Writing Research*. Taylor & Frances/Routledge; New York, NY: 2010. p. 166-180.
- Jenkins JR, Johnson E, Hileman J. When is reading also writing: Sources of individual differences on the new reading performance assessments. *Scientific Studies of Reading*. 2004; 8(2):25–151.
- Jajodia, A. Dynamic structural equation models of change.. In: Newsom, JT.; Jones, RN.; Hoker, SM., editors. *Longitudinal data analysis: A practical guide for researchers in aging, health, and social sciences*. Routledge; New York: 2012. p. 291-328.
- Jenkins JR, Johnson E, Hileman J. When is reading also writing: Sources of individual differences on the new reading performance assessments. *Scientific Studies of Reading*. 2004; 8(2):25–151.
- Juel C, Griffith P, Gough P. Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology*. 1986; 78:243–255.
- Juel C. Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*. 1988; 80:437–447.

- 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(3):281–300.
- Lerkkanen M, Rasku-Puttonen H, Aunola K, Nurmi J. The developmental dynamics of literacy skills during the first grade. *Educational Psychology*. 2004; 24(6):793–810.
- Little RJA, Rubin DB. The analysis of social science data with missing values. *Sociological Methods and Research*. 1989; 18:292–326.
- Loban, WD. The language of elementary school children. National Council of Teachers of English; Urbana, IL: 1963.
- McArdle JJ. Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*. 2009; 60:577–603.
- McArdle JJ, Grimm KJ. Five steps in latent curve and latent change score modeling with longitudinal data. *Longitudinal Research with Latent Variables*. 2010:245–273.
- Mehta P, Foorman B, Branum-Martin L, Taylor W. 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.
- Muthén, L.; Muthén, B. Mplus User's Guide. Sixth Edition.. Muthén & Muthén; Los Angeles, CA: 1998-2013.
- Nelson NW, Van Meter AM. Measuring written language ability in narrative samples. *Reading & Writing Quarterly*. 2007; 23:287–309.
- Pugh, KR.; Frost, SJ.; Sandak, R.; Gillis, M.; Moore, D.; Jenner, AR.; Mencl, WE. What does reading have to tell us about writing? Preliminary question and methodological challenges in examining the neurobiological foundations of writing and writing disabilities.. In: Charles, A.; MacArthur; Graham, Steve; Fitzgerald, Jill, editors. *Handbook of writing research*. Guilford Press; New York, NY: 2006. p. 433-448.
- Puranik C, Lonigan C. From scribbles to scrabble: Preschool children's developing knowledge of written language. *Reading and Writing*. 2010; 24:567–589. [PubMed: 22448101]
- Rode SS. Development of phrase and clause boundary reading in children. *Reading Research Quarterly*. 1975; 10(1):124–142.
- Shanahan T. Nature of reading-writing relations: An exploratory multivariate analysis. *Journal of Educational Psychology*. 1984; 76:466–477.
- Shanahan, T. Relations among oral language, reading, and writing development.. In: MacArthur, Charles A.; Graham, Steve; Fitzgerald, Jill, editors. *Handbook of writing research*. Guilford Press; New York, NY: 2006. p. 171-183.
- Shanahan T, Lomax RG. An analysis and comparison of theoretical models of the reading–writing relationship. *Journal of Educational Psychology*. 1986; 78(2):78, 116–123.
- Shanahan T, Lomax RG. A Developmental Comparison of Three Theoretical Models of the Reading-Writing Relationship. *Research in the Teaching of English*. 1988; 22(2):196–212.
- Sprenger-Charolles L, Siegel LS, Bechennec D, Serniclaes W. Development of phonological and orthographic processing in reading aloud, in silent reading, and in spelling: A four-year longitudinal study. *Journal of Experimental Child Psychology*. 2003; 84:194–217. [PubMed: 12706384]
- Thorndike, RL.; Hagen, E.; Sattler, J. *Stanford-Binet Intelligence Scales*. 4th edition.. Riverside; Itasca, IL: 1986.
- The Psychological Corporation. *Wechsler Abbreviated Scale of Intelligence*. Psychological Corp.; San Antonio, TX: 1999.
- Tierney, RJ.; Shanahan, T. Research on the reading–writing relationship: Interactions, transactions, and outcomes.. In: Barr, Rebecca; Kamil, Michael L.; Mosenthal, Peter B.; Pearson, P. David, editors. *Handbook of reading research*. Vol. 2. Lawrence Erlbaum Associates; Hillsdale, NJ, England: 1996. p. 246-280.
- Torrance, M.; Galbraith, D. The processing demands of writing.. In: MacArthur, Charles A.; Graham, Steve; Fitzgerald, Jill, editors. *Handbook of writing research*. Guilford Press; New York, NY: 2006. p. 67-80.

- Torgesen, JK.; Wagner, RK.; Rashotte, CA. Test of word reading efficiency. PRO-ED.; Austin, TX: 1999.
- Vaessen A, Blomert L. The cognitive linkage and divergence of spelling and reading development. *Scientific Studies of Reading*. 2013; 17:89–108.
- Vellutino FR, Tunmer WE, Jaccard J, Chen R, Scanlon DM. Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading*. 2004; 11(1):3–32.
- Wagner RK, Puranik CS, Foorman B, Foster E, Wilson LG, Tschinkel E, Kantor PT. Modeling the development of written language. *Reading & Writing: An Interdisciplinary Journal*. 2010; 24:203–220.
- Wagner, RK.; Torgesen, JK.; Rashotte, CA.; Pearson, N. The test of silent reading efficiency and comprehension (TOSREC). PRO-Ed.; Austin, TX: 2010.
- Weiser B, Mathes P. Using encoding instruction to improve the reading and spelling performance of elementary students at risk for literacy difficulties: A best-evidence synthesis. *Review of Educational Research*. 2011; 81:170–200.
- Wilkinson, GS. WRAT3 Wide range achievement test: Administration manual. Wide Range, Inc.; Wilmington, DE: 1993.
- Whitaker D, Berninger V, Johnston J, Swanson L. Intraindividual differences in levels of language in intermediate grade writers: Implications for the translating process. *Learning and Individual Differences*. 1994; 6:107–130.
- Woodcock, R. Woodcock reading mastery test—revised. American Guidance Service; Circle Pines, MN: 1987.
- Woodcock, RW.; Johnson, MB. Woodcock–Johnson tests of achievement. Riverside Publishing; Itasca, IL: 1990.
- Woodcock, RW.; McGrew, KS.; Mather, N. WJIII tests of cognitive abilities and achievement. Riverside Publishing; Itasca, IL: 2001.

Figure 1.

Unstandardized Estimates for the Word Level Reading-to-Writing Model. *Note.* a) auto-proportions b) auto-regressions c) status to change d) change to change.

* $p < 0.05$. ** $p < 0.001$.

Figure 2.
Unstandardized Estimates for the Sentence Level Bidirectional Model.
* $p < 0.05$. ** $p < 0.001$.

Figure 3.

Unstandardized Estimates for the Fluency Level Reading-to-Writing Model.

* $p < 0.05$. ** $p < 0.001$.

Figure 4.

Unstandardized Estimates for the Discourse Level Reading-to-Writing Model.

* $p < 0.05$. ** $p < 0.001$.

Table 1

Means, Standard Deviations, Minimum and Maximum Raw Scores and Z-Scores for the Reading and Writing Indicators

Variable Name	Raw Scores				Z-Scores			
	M	SD	Min	Max	M	SD	Min	Max
Grade 1 (n = 316)								
TOWRE NW D	14.03	8.67	0.00	34.00	0.00	1.00	-1.62	2.30
TOWRE NW C	15.63	8.78	0.00	40.00	0.00	1.00	-1.78	2.78
TOSREC Form A	18.52	10.88	0.00	42.00	0.00	1.00	-1.70	2.16
TOSREC Form O	17.48	10.37	0.00	44.00	0.00	1.00	-1.69	2.56
WJ Passage Comp.	463.53 ⁺	17.53	415.00	503.00	0.00	1.00	-2.77	2.25
WRMT P. Comp.	468.89 ⁺	15.46	427.00	500.00	0.00	1.00	-2.71	2.01
ORF 1	56.53	37.14	0.00	176.00	0.00	1.00	-2.00	3.00
ORF 2	52.85	35.06	3.00	142.00	0.00	1.00	-1.00	3.00
WRAT Spelling	21.53	3.45	11.00	30.00	0.00	1.00	-3.06	2.46
WIAT Spelling	17.30	5.02	4.00	31.00	0.00	1.00	-2.65	2.73
WIAT Sentences	0.95	1.12	0.00	5.00	0.00	1.00	-1.64	2.44
No. of Words	44.53	20.18	9.00	97.00	0.00	1.00	-1.76	2.60
WUR	3.78	0.88	0.00	6.00	0.00	1.00	-4.30	2.52
Grade 2 (n = 270)								
TOWRE NW B	25.75	11.35	1.00	53.00	1.35	1.31	-1.50	4.49
TOWRE NW A	26.08	10.54	5.00	53.00	1.19	1.20	-1.21	4.26
TOSREC Form A	25.59	9.41	0.00	50.00	0.65	0.87	-1.70	2.89
TOSREC Form O	26.00	9.38	0.00	50.00	0.82	0.90	-1.69	3.14
WJ Passage Comp.	483.81 ⁺	13.73	445.00	515.00	1.16	0.78	-1.06	2.94
WRMT P. Comp.	484.76 ⁺	11.38	429.00	516.00	1.03	0.74	-2.58	3.05
ORF 1	108.39	37.60	11.00	207.00	1.58	1.07	-1.00	4.00
ORF 2	100.25	38.10	9.00	206.00	1.32	1.06	-1.00	4.00
WRAT Spelling	24.75	3.91	15.00	36.00	0.93	1.14	-1.89	4.20
WIAT Spelling	23.61	5.52	11.00	38.00	1.26	1.10	-1.26	4.13
WIAT Sentences	2.30	1.74	0.00	6.00	1.36	1.14	-0.83	4.31
No. of Words	54.93	26.60	9.00	124.50	0.52	1.32	-1.76	3.96
WUR	3.83	0.91	0.00	6.00	0.06	1.03	-4.30	2.50
Grade 3 (n = 260)								
TOWRE NW D	30.34	10.31	7.00	55.00	1.82	1.27	-0.93	4.49
TOWRE NW C	29.80	11.02	6.00	53.00	1.68	1.17	-0.98	4.48
TOSREC Form A	28.57	8.37	7.00	51.00	0.92	0.77	-1.06	2.99
TOSREC Form O	27.96	9.08	0.00	54.00	1.01	0.88	-1.69	3.52
WJ Passage Comp.	494.20 ⁺	12.17	461.00	521.00	1.75	0.69	-0.14	3.28
WRMT P. Comp.	494.27 ⁺	11.68	430.00	522.00	1.64	0.76	-2.52	3.43

Variable Name	Raw Scores				Z-Scores			
	M	SD	Min	Max	M	SD	Min	Max
ORF 1	113.66	37.07	22.00	235.00	1.73	1.06	-1.00	5.00
ORF 2	102.79	38.42	8.00	217.00	1.39	1.07	-1.00	5.00
WRAT Spelling	28.76	3.73	21.00	43.00	2.10	1.08	-0.15	6.23
WIAT Spelling	28.30	5.75	16.00	44.00	2.19	1.15	-0.26	5.32
WIAT Sentences	3.81	2.21	0.00	10.00	0.41	0.76	-1.29	2.21
No. of Words	93.15	36.19	17.00	191.50	2.41	1.79	-1.36	7.28
WUR	3.94	0.81	0.00	6.00	-.06	0.86	-2.60	2.00
Grade 4 (<i>n</i> = 219)								
TOWRE NW B	34.69	10.50	11.00	58.00	2.17	1.19	-0.53	4.83
TOWRE NW A	35.27	11.18	9.00	62.00	2.45	1.29	-0.58	5.53
TOSREC Form A	29.97	9.62	6.00	60.00	1.05	0.88	-1.15	3.18
TOSREC Form O	29.71	9.39	3.00	58.00	1.18	0.91	-1.40	3.91
WJ Passage Comp.	499.92 ⁺	10.80	454.00	528.00	2.08	0.62	-0.54	3.68
WRMT P. Comp.	500.31 ⁺	11.00	461.00	524.00	2.03	0.71	-0.51	3.56
ORF 1	112.29	35.09	27.00	221.00	1.70	1.00	-1.00	5.00
ORF 2	118.55	30.58	24.00	205.00	1.83	0.85	-1.00	4.00
WRAT Spelling	30.57	4.70	18.00	43.00	2.62	1.36	-1.02	6.23
WIAT Spelling	31.72	6.10	16.00	45.00	2.88	1.22	-0.26	5.52
WIAT Sentences	5.02	2.47	0.00	11.00	3.64	2.21	-0.85	8.97
No. of Words	122.94	41.75	12.00	245.00	3.88	2.07	-1.61	9.94
WUR	3.33	1.02	0.00	5.50	-.78	0.66	-2.03	0.25

⁺ These values represent *w* scores rather than raw scores for passage comprehension measures for each year.

Table 2

Correlations among Variables for First Grade (Below Diagonal) and Second Grade (Above Diagonal)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. TOWRE 1	---	.94	.68	.69	.62	.54	.74	.77	.70	.64	.28	.12	.39
2. TOWRE 2	.92	---	.69	.71	.64	.54	.76	.79	.71	.66	.28	.11 ^{ns}	.43
3. TOSREA	.79	.79	---	.85	.76	.67	.78	.77	.63	.68	.38	.22	.43
4. TOSREO	.78	.78	.92	---	.76	.68	.80	.81	.64	.67	.38	.22	.49
5. WJPC	.73	.72	.84	.84	---	.77	.76	.71	.64	.71	.39	.20	.50
6. WRMT PC	.76	.75	.85	.85	.87	---	.66	.61	.54	.64	.37	.10 ^{ns}	.47
7. ORF 1	.84	.83	.91	.89	.82	.82	---	.94	.62	.64	.28	.18	.43
8. ORF 2	.82	.82	.91	.88	.82	.82	.96	---	.61	.62	.23	.19	.41
9. WRAT SP	.72	.72	.74	.74	.70	.74	.76	.75	---	.83	.39	.12	.49
10. WIAT SP	.72	.73	.79	.80	.78	.80	.78	.77	.81	---	.44	.14	.50
11. WIAT SEN	.29	.30	.35	.34	.38	.38	.35	.36	.32	.34	---	.09 ^{ns}	.36
12. No. of Words	.35	.33	.41	.41	.33	.36	.41	.40	.38	.36	.11	---	.20
13. WUR	.42	.42	.49	.51	.49	.59	.49	.44	.44	.54	.25	.31	---

Note. TOWRE 1= Test of Silent Reading Efficiency: Phonetic Decoding Efficiency Form D (grade 1) or Form B (grade 2), TOWRE 2 = Test of Silent Reading Efficiency: Phonetic Decoding Efficiency Form C (grade 1) or Form A (grade 2), TOSREC A = Test of Silent Reading Efficiency and Comprehension Form A, TOSREC O = Test of Silent Reading Efficiency and Comprehension Form O, WJ PC = Woodcock Johnson Passage Comprehension, WRMT PC = Woodcock Reading Mastery Test Passage Comprehension, ORF = Dynamic Indicators of Basic Early Literacy Skills Oral Reading Fluency, WRAT Sp = Wide Range Achievement Test Spelling, WIAT Sp = Wechsler Individual Achievement Test Spelling, WIAT SEN = Wechsler Individual Achievement Test Combining Sentences, WUR = Writes Upon Request. Correlations greater than .14 are significant at the 0.001 level. Correlations greater than .11 are significant at the 0.05 level.

^{ns} Correlations are not significant.

Table 3

Correlations among Variables for Third Grade (Below Diagonal) and Fourth Grade (Above Diagonal)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. TOWRE 1	---	.94	.63	.64	.50	.57	.74	.76	.70	.68	.40	.28	.38
2. TOWRE 2	.93	---	.62	.62	.49	.55	.74	.75	.70	.66	.40	.29	.34
3. TOSREC A	.68	.68	---	.84	.64	.68	.80	.80	.66	.64	.50	.37	.48
4. TOSREC O	.66	.65	.82	---	.62	.69	.79	.81	.64	.63	.54	.40	.41
5. WJ PC	.62	.62	.74	.66	---	.78	.64	.63	.53	.55	.49	.35	.39
6. WRMT PC	.55	.51	.67	.59	.74	---	.68	.67	.62	.59	.50	.30	.40
7. ORF 1	.78	.78	.83	.75	.69	.67	---	.92	.67	.68	.45	.38	.36
8. ORF 2	.79	.76	.82	.76	.70	.67	.93	---	.70	.71	.47	.40	.43
9. WRAT SP	.68	.69	.65	.58	.66	.60	.69	.69	---	.86	.43	.35	.39
10. WIAT SP	.67	.68	.66	.63	.66	.61	.73	.73	.84	---	.49	.37	.38
11. WIAT SEN	.47	.46	.55	.53	.58	.56	.56	.56	.47	.55	---	.31	.42
12. No. of Words	.21	.23	.34	.34	.36	.24	.30	.30	.27	.36	.29	---	.28
13. WUR	.37	.34	.47	.44	.48	.44	.48	.46	.39	.51	.43	.25	---

Note. TOWRE 1= Test of Silent Reading Efficiency: Phonetic Decoding Efficiency Form D (grade 3) or Form B (grade 4), TOWRE 2 = Test of Silent Reading Efficiency: Phonetic Decoding Efficiency Form C (grade 3) or Form A (grade 4), TOSREC A = Test of Silent Reading Efficiency and Comprehension Form A, TOSREC D = Test of Silent Reading Efficiency and Comprehension Form O, WJ PC = Woodcock Johnson Passage Comprehension, WRMT PC = Woodcock Reading Mastery Test Passage Comprehension, ORF = Dynamic Indicators of Basic Early Literacy Skills Oral Reading Fluency, WRAT Sp = Wide Range Achievement Test Spelling, WIAT Sp = Wechsler Individual Achievement Test Spelling, WIAT SEN = Wechsler Individual Achievement Test Combining Sentences, WUR = Writes Upon Request.

Correlations greater than .16 are significant at the 0.001 level. Correlations greater than .12 are significant at the 0.05 level.

nsCorrelations are not significant.

Table 4

Model Fit Indices

Model	χ^2	df	CFI	TLI	RMSEA	SRMR	χ^2 difference (vs. bidirectional)
Word Level							
Bidirectional	417.53	123	0.95	0.95	0.09	0.06	
R-to-W	418.70	125	0.95	0.95	0.09	0.06	1.17
W-to-R	481.41	124	0.94	0.94	0.10	0.09	63.88**
Sentence Level							
Bidirectional	126.52	63	0.98	0.98	0.06	0.07	
R-to-W	148.78	65	0.97	0.97	0.06	0.08	22.27**
W-to-R	223.34	65	0.94	0.94	0.09	0.18	96.82**
Text Level (Fluency)							
Bidirectional	290.06	63	0.92	0.92	0.12	0.09	
R-to-W	291.15	65	0.92	0.92	0.12	0.09	1.14
W-to-R	301.70	65	0.92	0.92	0.12	0.10	46.31**
Text Level (Discourse)							
Bidirectional	161.04	63	0.94	0.94	0.08	0.10	
R-to-W	165.19	65	0.94	0.94	0.08	0.11	1.09
W-to-R	207.25	65	0.92	0.92	0.09	0.13	11.64*

* $p < 0.05$ ** $p < 0.001$.

Table 5

Sample Statistics for the Estimated Change Scores for the Best Fitting Models

	Reading			Writing		
	1	2	3	1	2	3
Word						
M	1.27	0.48	0.53	1.02	1.00	0.57
SD	0.57	0.18	0.18	0.26	0.20	0.25
Min	-0.17	0.02	-0.11	0.42	0.54	-0.04
Max	3.12	1.08	1.21	2.05	1.66	1.47
Sentence						
M	0.67	0.23	0.13	1.16	1.36	1.06
SD	0.14	0.19	0.18	1.35	1.55	1.49
Min	-0.50	-0.31	-0.47	-2.90	-2.40	-4.09
Max	2.21	0.75	0.81	4.92	6.76	5.92
Text (Fluency)						
M	1.27	0.10	0.10	0.33	2.06	1.33
SD	0.44	0.19	0.12	1.07	1.40	1.27
Min	0.37	-0.45	-0.31	-2.57	-1.96	-2.77
Max	2.77	0.68	0.49	4.04	5.65	5.51
Text (Discourse)						
M	0.94	0.60	0.30	-0.01	-0.14	-0.43
SD	0.24	0.16	0.11	0.77	0.61	0.72
Min	0.05	-0.16	-0.03	-2.17	-2.38	-4.27
Max	1.82	1.41	0.77	2.08	1.69	1.30

Table 6

Sample Statistics for the Estimated Change Scores by Gender

	Reading						Writing					
	Males			Females			Males			Females		
	1	2	3	1	2	3	1	2	3	1	2	3
Word												
M	1.32	0.52	0.54	1.23	0.44	0.53	1.04	0.98	0.54	1.01	1.03	0.58
SD	0.55	0.20	0.13	0.58	0.18	0.23	0.29	0.22	0.23	0.24	0.19	0.27
Min	-0.18	0.03	0.19	-0.05	0.05	-0.23	0.35	0.50	-0.02	0.57	0.64	-0.01
Max	3.13	1.15	0.95	3.00	0.94	1.33	1.92	1.72	1.43	1.89	1.61	1.33
Sentence												
M	0.71	0.20	0.17	0.62	0.26	0.08	1.24	1.2	1.14	1.06	1.52	1.00
SD	0.38	0.20	0.15	0.38	0.18	0.21	1.39	1.54	1.47	1.31	1.57	1.51
Min	-0.46	-0.33	-0.29	-0.26	-0.18	-0.59	-2.07	-2.36	-2.47	-2.94	-2.37	-4.09
Max	2.21	0.71	0.68	1.69	0.76	0.91	4.93	5.08	5.92	4.24	6.79	5.08
Text (Fluency)												
M	1.32	0.15	0.20	1.37	0.07	0.09	0.39	1.59	1.37	0.56	2.20	1.60
SD	0.47	0.24	0.19	0.55	0.20	0.13	0.95	1.47	1.31	1.30	1.46	1.41
Min	-0.09	-0.58	-0.34	0.34	-0.51	-0.28	-1.81	-3.34	-2.78	-2.69	-1.43	-2.35
Max	2.67	1.13	0.85	3.51	0.68	0.44	3.82	5.64	5.23	4.21	5.55	6.68
Text (Discourse)												
M	0.99	0.63	0.33	0.97	0.63	0.27	0.02	-0.1	-0.48	0.18	1.15	1.77
SD	0.24	0.23	0.16	0.26	0.11	0.09	0.74	0.59	0.78	0.76	0.69	0.6
Min	0.14	-0.54	-0.21	0.10	0.39	0.09	-1.75	-1.62	-4.32	-2.08	-1.13	-0.13
Max	1.81	1.63	1.15	2.00	1.00	0.56	1.67	1.67	1.32	1.78	2.50	2.92