2011

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

Yusra Ahmed

THE FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

DEVELOPMENTAL RELATIONS BETWEEN READING AND WRITING AT THE WORD, SENTENCE AND TEXT LEVELS: A LATENT CHANGE SCORE ANALYSIS

By

YUSRA AHMED

A Thesis submitted to the
Department of Psychology
in partial fulfillment of the
requirements for the degree of
Master of Science

Degree Awarded:
Fall Semester, 2011

Yusra Ahmed defended this thesis on August 8th, 2011.

The members of the supervisory committee were:

Richard K. Wagner
Professor Directing Thesis

Chris Schatschneider
Committee Member

Jon Maner
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the thesis has been approved in accordance with university requirements.

ACKNOWLEDGEMENTS

I would like to express my deepest thanks to my major professor Dr. Rick Wagner for his invaluable support, time and attention throughout my graduate career. His guidance and patience throughout have served as an inspiration for me. I would also like to thank my committee members, Dr. Chris Schatschneider and Dr. Jon Maner for their time and feedback on this study. My thanks are also extended to the researchers at the Wagner Lab (past and present) for timely data collection and management.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Purpose – The relations between reading and writing have been studied extensively but the exact nature of their interrelation is not known. This study applied new advances in change score modeling to investigate longitudinal developmental relations between reading and writing skills at the word, sentence and text levels. Dynamic models were used to compare unidirectional pathways (reading-to-writing and writing-to-reading) and bidirectional pathways in a test of nested models. **Method -** Participants included 316 boys and girls who were assessed annually in grades 1 through 3. Measures of reading included pseudo word decoding, sentence reading efficiency measures and passage comprehension. Measures of writing included spelling, a sentence combining task and computational indices of linguistic features of a writing prompt. **Results** – The changes in reading and writing were characterized by improvements between years. The reading-to-writing model fit the data well, where changes in reading and writing were a function of a) status, b) growth or c) both in reading or writing. At the word level, high status in grade 1 decoding predicted an improvement on spelling between grades 2-3, and the improvement in decoding between grades 1-2 predicted an improvement in spelling between grades 2-3. At the sentence level, high status in reading predicted an improvement in writing across the years. At the text level, high status in grade 2 reading predicted an improvement in writing between grades 2-3. **Discussion** – findings suggest that a reading-to-writing model better describes the data than does a bidirectional model, and this relation holds across levels of language. Secondly, changes in writing are predicted by high achievement status in reading at all levels of language, and this effect was significant at both stages of development at the word and sentence levels, and at the later stage of development at the text level (between grades 2 and 3). Thirdly, change in spelling was predicted by change decoding between grades 1-2. Thus, our results show that Mathew effects are characteristic of reading-writing development in that acquisition of writing skills is facilitated for good readers.

INTRODUCTION

Although historically most research and pedagogy has separated reading and writing instruction (Shanahan, 2006), the relations between reading and writing (i.e. literacy skills) have been studied widely 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; Langer & Filhan, 2000; Abbott & Berninger, 1993; Tierney & Shanahan, 1991; Juel et al., 1986; Juel, 1988, 1983; Loban, 1963; Shanahan, 1984), and interventions that have focused on transfer of skills show that it's possible to teach reading so that it improves writing (Shanahan, 2006) and writing so that it improves reading (Tierney & Shanahan, 1991). However, the exact nature of the interrelation is not known.

Earlier studies that examined relations between reading and writing were correlational in nature. In a study of more than 600 students in grades 1 – 6, Abbott and Berninger (1993) found that correlations between passage comprehension and a measure of compositional fluency (measured as number of words and clauses written) ranged from .08 to .34. Correlations between passage comprehension and a holistic measure of writing were higher, and ranged from .22 to .54. Shanahan (1984) examined the relation among four reading measures (vocabulary, phonics, comprehension, and sentence completion) and eight writing measures (vocabulary diversity, average t-unit length, number of episodes, number of story categories, number of information units, spelling accuracy, phonemic accuracy, and orthographic accuracy) for students in grades 2 and 5. His results showed that correlations between the reading and writing measures ranged from .14 (vocabulary and writing episodes) to .68 (phonics and orthographic accuracy).

Correlational studies provided evidence for a general association between reading and writing by showing that the shared variability between reading and writing exceeded 50%. Later studies included structural equation modeling and were longitudinal in nature (e.g., Metha et al., 2005; Lerkkanen et al., 2004; Shanahan & Lomax, 1986, 1988). Structural equation models include various measures to form reading and writing latent factors (i.e., factors that do not include measurement error). Because of the use of several indicators in later studies, estimates of shared variability between reading and writing at the word level have ranged from 77-85% and about 65% for text level variables (Berninger et al., 2002).

Cross-lagged structural equation models have also been used to examine the directionality of influences. In cross-lagged models, each common factor in the model influences itself over time (i.e., stability parameters) and each factor crosses over to influence the other factor at subsequent times. Shanahan and Lomax's (1986) study is one of the earliest examples of SEM techniques used to examine bidirectional reading-writing relations. Their study is similar to the research described in the current study in that they studied the relation between reading and writing across grades, though the current research analyzed three different grade levels rather than the two included in their study. Their analysis included three latent variables for reading (word analysis, vocabulary, and comprehension) and four for writing (spelling, vocabulary diversity, syntax, and story structure). Their bidirectional model was an interactive model in which writing factors influenced the reading factors and vice versa. In their unidirectional models, reading factors influenced writing factors and writing factors influenced reading factors, although both relations were not estimated simultaneously. Shanahan and Lomax compared the models both within each grade and across grade 2 and grade 5. They found that the model that represented an interaction between reading and writing factors provided the best fit to the data in both grades.

Lerkkanen, Rasku-Puttonen, Aunola & Nurmi (2004) conducted another study that examined the bidirectional relation 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 less spelling errors). Their analyses included one latent variable for reading (comprised of word reading and reading comprehension) and one latent variable for writing (comprised of spelling and writing fluency, which was measured as writing words, sentences or a story about a given picture). 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 the relations between reading and writing is bidirectional (Shanahan & Lomax, 1986; Lerkkanen et al., 2004; Abbott et al., 2010) where both reading and writing are uniquely predictive of one another, other studies suggest that the relation is unidirectional, with some studies finding that writing influences reading (e.g., Caravolas et al., 2001; Cataldo et al., 1988; Shanahan & Lomax, 1988; Berninger et al., 2002) and other studies finding that reading influences writing (Shanahan & Lomax, 1986; Aarnoutse et al., 2004; Babagayigit & Stainthorp, 2011; Sprenger-Charolles et al., 2003; Berninger et al., 2002). Evidence is mixed in terms of the functional form of relation between reading and writing, partly because of the number and type of indicators used to represent each construct. For example, of the studies mentioned above Lerkkanen et al. (2004) defined reading using a composite of word reading and reading comprehension measures as well as a composite of spelling and writing. On the other hand, Shanahan and Lomax (1986) included separate latent factors for reading (word analysis, vocabulary, and comprehension) and writing (spelling, vocabulary diversity, syntax, and story structure) in a single model. Thus, studies have differed in the use of cumulative vs. componential measures. An implicit assumption in the components based studies is that the measures chosen adequately produce estimates of overall reading or writing performance. However, including many components in one model makes them less parsimonious. Research that has focused on cumulative measures is based on the assumption that reading comprehension and composition are culminating activities that involve the use of multiple skills (e.g., word recognition, syntactic knowledge) and thus are more likely to capture many of the elements shared by the two processes than instruments that measure component skills in isolation. Although component based studies provide a fine grained examination of literacy processes, these studies have not established a symmetrical relation between components of reading and writing.

Levels of Language Approach

Recent studies have analyzed separate components of reading and writing based on levels of language. The *levels of language* approach is based on the size of the written code (e.g., word, sentence and passage). These levels are typically not related in a one-to-one fashion for reading and writing (Whitaker, et al., 1994; Berninger et al., 2002; Abbott et al., 2010), but they provide a way to compare components of each that are similar. 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 (Whitaker, Berninger, Johnston, & Swanson, 1994; Abbott & Berninger, 1993), 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 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 which 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. Although it is possible that like readers, writers rely on phonology for encoding words (Dell, 1986), 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 (Foorman et al., 2011; Abbott et al., 2010, 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 constrains in oral language as well as written language (e.g., Rode, 1976; Bates & Goodman, 1999), even at the preschool level (Puranik & Lonigan, 2010). Although syntactic knowledge can be derived from oral language most of the syntactic structures found in written language are learned through reading. 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 sentence begins 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. Text format refers to the syntax of larger texts and their structural organization (Shanahan, 2006). Older studies, such as Juel, Griffith & Gough (1986) have found that passage comprehension and writing composition have a weak relation. However, these studies have used holistic measures of writing ability which have many limitations. More recent studies have used compositional fluency and quality measures and show that the correlation between passage comprehension and text composition are high for both children and adults (Berninger et al., 2002) and that reading comprehension and composition are mutually predictive over time (Abbott et al., 2010). Because texts consist of clauses and sentences, readers apply a series of inferences and construct prepositions based on the information provided by the sentences (Foorman et al., 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 and it entails translation of ideas into writing as well as transcription (handwriting and spelling) skills (Hayes & Berninger, 2010). As mentioned previously, text level reading and writing are often considered outcome processes that encompass word and sentence level processes. Thus, studies that have separated text level literacy development from word and sentence level are scarce.

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 the simple view of writing holds that writing is a product of word level writing (spelling) and text level writing (composition) (Berninger, 2002b). 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 that the relation between reading and writing at the word and text levels is 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. Abbott et al. claimed that the 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 single observed indicators as measures of reading and writing ability. The use of single indicators assumes that the constructs were measured without error, and the stability parameters could largely be inflated due to common method variance. In addition, they used a measure of composition, the WIAT II Written Expression, which combines word, sentence and text level writing and may have been a poor indicator of text-level writing. Finally, their study did not include a sentence level model.

The present study extends on previous research in three ways: a) the use of several time points, b) the use multiple indicators to represent reading and writing across levels of language, and c) the use of dynamic models. The present study applied a different change concept than previous studies, whereby latent change score models were favored over cross-lagged models to formulate different hypotheses of change over time. The use of dynamic modeling allows for the separation of variability that is due to achievement status from variability that is due to change. This is particularly important because previous research has assumed that a significant path between a reading and writing variable indicates a strong relation. For example, a significant pathway from spelling to reading indicated that spelling is a leading indicator of reading. However, it is possible to examine whether a variable influences another variable by modeling the effect that improvement in one variable has on the improvement on another variable. Thus, this study examined the effect of changes in reading and writing on subsequent changes in reading or writing, at the word, sentence and text levels. The next section provides a description of the models used in this study. For a comprehensive review of change models see McArdle (2009).

Latent Change Score Modeling

Latent change scores (LCS) have been applied to various areas of research, including human development, social and personality psychology, psychopathology, cognitive psychology and neuropsychology to examine dynamics among psychological processes (Ferrer & McArdle, 2007). LCS models take advantage of characteristics of structural equation models and latent growth curve models by accounting for the covariance structure as well as the mean structure. Similar to cross-lagged structural equation models, change score models include stability parameters as well as cross regressions, and both models are based on time-sequence to determine the flow of one variable from another. Although both models describe processes that unfold over time and involve multiple interrelated parts, change score analysis is a more appropriate developmental method because it separates change (which could be an increase or decrease in reading or writing) from its common factor (McArdle, 2009; McArdle & Grimm, 2010). The question that a cross-lagged model answers is: *does reading influence writing*? Conversely, the question that a latent change score models answers is: *does an improvement in reading lead to an improvement in writing?*

Similar to latent growth curve models, LCS models describe both group average change overtime and individual differences around these changes. Unlike growth curve models, the growth estimates are not based on slopes (or rate of change) but on differences between two scores. Thus, an important distinction is that change score models are not based on residualized change. The changes are essentially the subtraction between the common factor at a time point (t) and the previous time point (t - 1) (McArdle, 2009). Furthermore, the estimated score at each time point is the status in achievement at that time point, where the first score is the initial score. Because the common factors are free of unique variance such as error variance and systematic variance, the latent change scores are free of measurement error.

A bivariate model that includes bidirectional pathways is illustrated in **Figure 1** where the latent change scores are represented by the symbol (Δ). The stability parameters (autoregressors) and the pathway from the latent change score to its common factor are fixed at 1 to meet model identification (McArdle, 2009; McArdle & Grimm, 2010). The triangle represents a constant that is used to estimate latent means. The correlation between the latent factors at Year 1 is included to account for causes not explained by the model. As mentioned above, the first latent factors (Year 1 reading and Year 1 writing) are the initial reading and writing scores, and the first changes (Δ_{r1} and Δ_{w1}) are the initial change scores. Univariate parameters are regression coefficients of the change score regressed on its common factor at an earlier time point. Multivariate paths include the change scores regressed on the other common factor at an earlier time point.

Given the layout of Figure 1, additional pathways can be added to estimate causal effects of the initial change scores on the second change scores. Specifically, autoregressors and cross-regressions can be added between the change scores as depicted in **Figure 2**. The additional impact of these pathways allows the change in one variable to impact the change in the other.

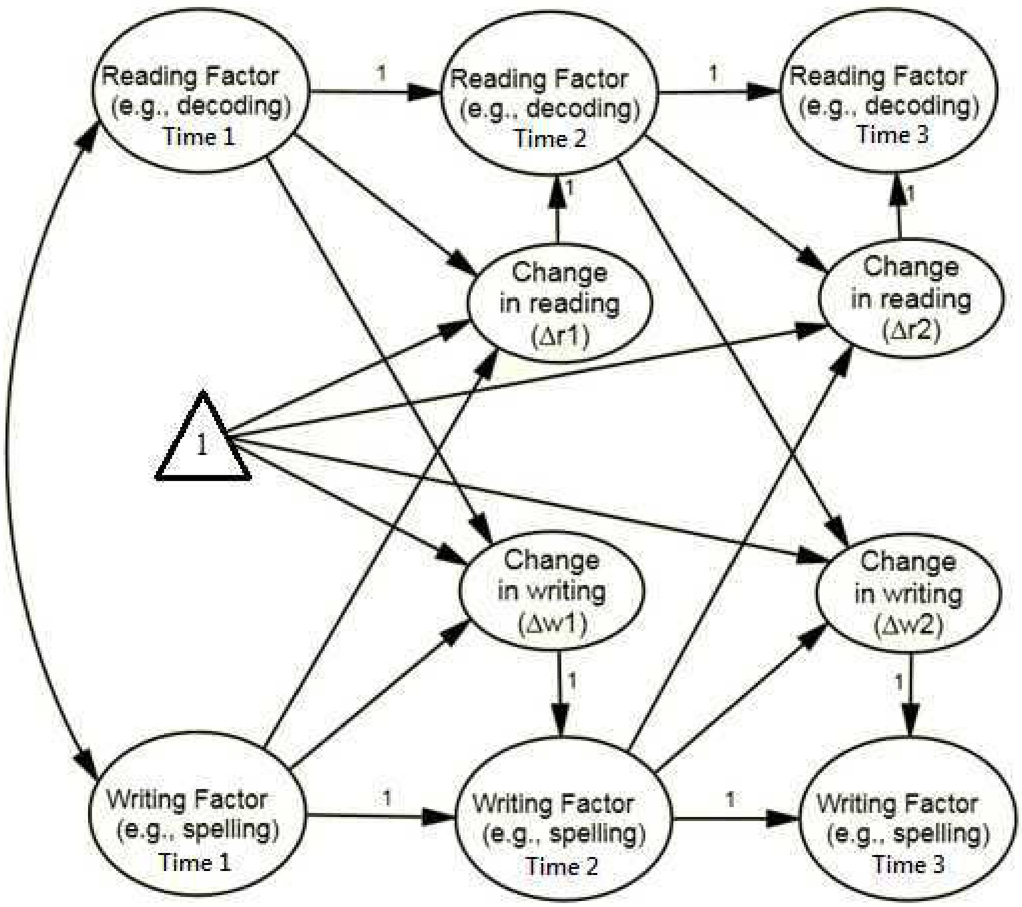

Figure 1. Structural Portion of the Latent Change Score Model

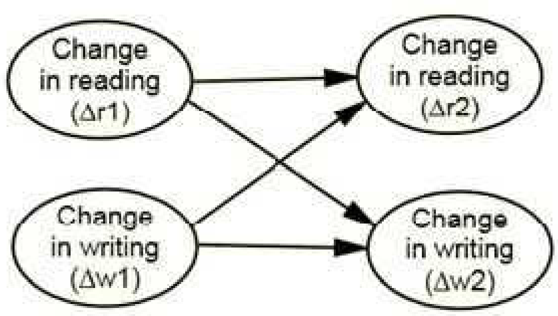

Figure 2. Bivariate Relations between Change Scores

It's important to note that initial change scores (Δ_{r1} and Δ_{w1} in Figures 1-2) are different from subsequent change scores (Δ_{r2} and Δ_{w2} in Figures 1-2) because of the additional parameters that predict the second change scores. Thus, all change scores are predicted by the reading and writing latent factors at the previous year, but the second change scores are also predicted by the initial change scores.

The equations used for the specification of the full model are described next. Following classical factor-analysis theory, for each person *i* at time *t* (with subscripts for individuals suppressed), the observed indicator (R for reading and W for writing) is a function of a) its initial score (r_0 or w_0) b) a trajectory component which consists of the accumulation of latent change scores up to that point ($\sum\Delta_r$ or $\sum\Delta_w$) and a constant (k) and c) the residual errors (Ferrer & McArdle, 2010; McArdle, 2009):

$$R_t = r_0 + (\sum\Delta_r k) + e_{rt} \quad (1a)$$

$$W_t = w_0 + (\sum\Delta_w k) + e_{wt} \quad (1b)$$

The latent change score, Δ_{rt} or $\Delta_{wt,}$ is defined as a function of two components, as shown in equations 2a and 2b. The first component includes (β), a self-feedback coefficient between the LCS and its common factor at a previous time point ($r_{t-1} \rightarrow \Delta_{rt}$ and $w_{t-1} \rightarrow \Delta_{wt}$). The second component is the coupling parameter between the LCS and the common factor of the other variable at a previous time point ($w_{t-1} \rightarrow \Delta_{rt}$ and $r_{t-1} \rightarrow \Delta_{wt}$). For Δ_{r1} in Figures 1-2, the first component is $\Delta_{r1} \rightarrow$ Reading Factor Time 1 and the second component is $\Delta_{r1} \rightarrow$ Writing Factor Time 1. The equations for the LCS are presented below:

$$\Delta_{r1} = (\beta_r * r_{t-1}) + (\gamma_r * w_{t-1}) \quad (2a)$$

$$\Delta_{w1} = (\beta_w * w_{t-1}) + (\gamma_w * r_{t-1}) \quad (2b)$$

The second change scores in the reading and writing models (Δ_{r2} and Δ_{w2} in Figure 2) have two additional components, as shown in equations 3a and 3b. The first component is the autoregressor of the change score regressed on itself at a previous time point ($\Delta_{r1} \rightarrow \Delta_{r2}$ and Δ_{w1}

$\rightarrow \Delta_{w2}$) and the second component is a causal pathway between the change score of one factor and the change score of the second factor ($\Delta_{w1} \rightarrow \Delta_{r2}$ and $\Delta_{r1} \rightarrow \Delta_{w2}$).

$$\Delta_{r2} = (\beta_r * r_{t-1}) + (\gamma_r * w_{t-1}) + (\alpha_r * \Delta_{r1}) + (\theta_r * \Delta_{w1}) \tag{3a}$$

$$\Delta_{w2} = (\beta_w * w_{t-1}) + (\gamma_w * r_{t-1}) + (\alpha_w * \Delta_{w1}) + (\theta_w * \Delta_{r1}) \tag{3b}$$

For the Reading-to-Writing model, the equations for the changes in reading did not include regressions on writing factors, as shown below. The change scores for writing remained the same as the bidirectional (full) model (equations 2b and 3b).

$$\Delta_{r1} = (\beta_r * r_{t-1}) \tag{4a}$$

$$\Delta_{r2} = (\beta_r * r_{t-1}) + (\alpha_r * \Delta_{r1}) \tag{4b}$$

For the Writing-to-Reading model, the equations for the changes in writing did not include regressions on reading factors, as shown below. The change scores for reading remained the same as the bidirectional model (equations 2a and 3a).

$$\Delta_{w1} = (\beta_w * w_{t-1}) \tag{5a}$$

$$\Delta_{w2} = (\beta_w * w_{t-1}) + (\alpha_w * \Delta_{w1}) \tag{5b}$$

METHOD

Procedures

The data used in this study was obtained from a four year longitudinal study assessing various aspects of literacy including decoding, fluency, vocabulary, reading comprehension, listening comprehension, working memory and writing. 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 was gathered by trained testers during thirty to sixty minute sessions in a quiet room designated for testing at the schools. The test battery was scored in a lab by two or more raters and discrepancies in the scoring were resolved by an additional rater. Because the last wave of data was being collected at the time of this study, data was used for years one through three.

Participants

In the first year of data collection 318 consent forms were returned. However, data were missing for 2 participants and they were dropped from the study. The first year sample therefore consisted of 316 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, 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.

Measures

Reading Measures

Decoding Measures. The Woodcock Reading Mastery Tests-Revised (WRMT-R; Woodcock, 1987): Word Attack subtest was used to assess accuracy for decoding non-words. The Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999): Phonetic Decoding Efficiency (PDE) subtest was also used to assess pseudo-word reading fluency and accuracy. Both subtests were used to form a word level decoding latent factor. The WRMT-R Word Attack subtest consist of a list of non-words that are read out loud by the participant. The lists start off with letters and become increasingly more difficult to include complex non-words. Testing is discontinued after six consecutive incorrect items. The median reliability is reported to be .87 for Word Attack (Woodcock, McGrew, & Mather, 2001). 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 Measures. Two forms of the *Test of Silent Reading Efficiency and Comprehension* (*TOSREC, forms A and D*; 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.

Reading Comprehension Measures. 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.

Writing Measures

Spelling Measures. 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 Measures. 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. A writing composition task was also administered. Participants were asked to write a passage on a topic provided by the tester. Students were instructed to scratch out any mistakes and were not allowed to use erasers. The task was administered in groups and lasted 10 minutes. The passages for years 1 and 2 required expository writing and the passage for year 3 required narrative writing. The topics were as follows: choosing a pet for the classroom (year 1), favorite subject (year 2), a day off from school (year 3). The writing samples were transcribed into a computer database by two trained coders. In order to submit the samples to Coh-Metrix (described below) the coders also corrected the samples. Samples were corrected once for spelling and punctuation using a hard criterion (i.e., words were corrected individually for spelling errors regardless of the context, and run-on sentences were broken down into separate sentences). In addition, the samples were completely corrected using the soft criterion: corrections were made for spelling based on context (e.g., correcting *there* for *their*), punctuation, grammar, usage, and syntax (see **Appendix A** for examples of original and corrected transcripts). The samples that were corrected only for spelling and punctuation using the hard criterion were used for several reasons: (a) developing readers make many spelling errors which make their original samples illegible, and (b) the samples that were completely corrected do not stay true to the child's writing ability. Accuracy of writing was not reflected in the corrected samples because of the elimination of spelling errors. However, as mentioned above spelling ability was measured separately. Data on compositional fluency and complexity were obtained from Coh-Metrix. Compositional fluency refers to how much writing was done and complexity refers to the density of writing and length of sentences (Berninger et al., 2002; Wagner et al., 2010).

Coh-Metrix Measures. The transcribed samples were analyzed using Coh-Metrix (McNamara et al., 2005; Graesser et al., 2004). Coh-Metrix is a computer scoring system that analyzes over 50 measures of coherence, cohesion, language, and readability of texts. **Appendix B** contains the list of variables provided by Coh-Metrix. In the present study, the variables were broadly grouped into the following categories: a) syntactic, b) semantic, c) compositional fluency, d) frequency, e) readability and f) situation model. *Syntactic* measures provide information on pronouns, noun phrases, verb and noun constituents, connectives, type-token ratio, and number of words before the main verb. Connectives are words such as *so* and *because* that are used to connect clauses. Causal, logical, additive and temporal connectives indicate cohesion and logical ordering of ideas. Type-token ratio is the ratio of unique words to the number of times each word is used.

Semantic measures provide information on nouns, word stems, anaphors, content word overlap, Latent Semantic Analysis (LSA), concreteness, and hypernyms. Anaphors are words (such as pronouns) used to avoid repetition (e.g., *she* refers to a person that was previously described in the text). LSA refers to how conceptually similar each sentence is to every other sentence in the text. Concreteness refers to the level of imaginability of a word, or the extent to which words are not abstract. Concrete words have more distinctive features and can be easily pictured in the mind. Hypernym is also a measure of concreteness, and refers to the conceptual taxonomic level of a word (for example, *chair* has 7 hypernym levels: seat -> furniture -> furnishings -> instrumentality -> artifact -> object -> entity).

Compositional fluency measures include the number of paragraphs, sentences and words, as well as their average length and the frequencies of content words. *Frequency* indices provide information on the frequency of content words, including several transformations of the raw frequency score. Content words are nouns, adverbs, adjectives, main verbs, and other categories with rich conceptual content. *Readability indices* are related to fluency and include two traditional indices used to assess difficulty of text: Flesch Reading Ease Score and Flesch-

Kincaid Grade Level. Finally, *situation model* indices describe what the text is about, including causality of events and actions, intentionality of performing actions, tenses of actions and spatial information.

Because Coh-Metrix hasn't been widely used to study the development of writing in primary grade children (Puranik et al., 2010) the variables used in the present study were determined in an exploratory manner described below. Out of the 56 variables, 3 were used in the present study: total number of words, total number of sentences and average sentence length (or average number of words per sentence). Nelson and Van Meter (2007) report that total word productivity is a robust measure of developmental growth in writing. Therefore, indicators for a paragraph level factor included total number of words and total number of sentences. Average words per sentence was used as an indicator for a latent sentence level factor, along with the WIAT-II Combining Sentences task.

RESULTS

Data Issues and Descriptive Statistics

Prior to analysis, data were screened for outliers, missing values, and normality. As previously noted, data were missing for 2 participants for year 1 and were therefore dropped from the study. Furthermore, the attrition rate was 14.5% (n= 46) at Time 2 and 3.7% (n =10) at Time 3. Reasons for not being included in years 2 and 3 were moving out of the area, no longer wishing to participate, or unable to be contacted. Additionally, data were missing for 1 participant for each of the following: T1 TOWRE, T1 TOSREC Form D, T1 WJ Passage Comprehension, T2 WIAT Spelling and T3 WIAT Spelling. Full-information maximum-likelihood estimation (FIML) was used in Mplus 6 (Muthén & Muthén, 1998-2007) 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 brought to the fence using the median +/- two interquartile range as criterion (Tabachnik & Fidell, 2007). The following standardized variables were adjusted: T1 WJ Passage Comprehension (2 values), T1 WRMT Passage Comprehension (1 value), T1 WIAT Spelling (1 value), T2 WJ Passage Comprehension (3 values), T2 WIAT Spelling (1 value), T2 WIAT Combining Sentences (10 values), T3 TOSREC Form A (3 values), and T3 WJ Passage Comprehension (4 values). The following Coh-Metrix variables were also adjusted[1]: T1 average sentence length (23 values), T1 number of sentences (10 values), T1 number of words (5 values) T2 average sentence length (14 values), T2 number of sentences (16 values), T2 number of words (21 values) T3 average sentence length (13 values), T3 number of sentences (2 values), T3 number of words (2 values). Visual inspection of scatter plots revealed no bivariate outliers.

Evaluation of skewness and kurtosis statistics revealed mild departure from normality for T2 WRMT Passage Comprehension (kurtosis = 2.49) and T3 WRMT Passage Comprehension (kurtosis = 2.92), likely due to few data points hitting the floor on these variables (n=3 for year 2; n=1 for year 3). The remaining skewness and kurtosis values fell within an acceptable range indicating that the data were normally distributed.

[1] Over 130 Coh-Metrix variables were adjusted across time points, but only the ones used in the final analyses are reported here.

As a final step each child's raw scores (or *w* scores[2]) 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 score. Z scores are useful for comparing values of variables that are measured on different scales. **Table 1** contains descriptive statistics for all measures. As noted in the table, means on all measures increase from year. **Tables 2-4** contain correlations reported by year. Examination of the correlations between the standardized measures of reading and the Coh-Metrix measures reveals a weaker relation than between standardized measures of reading and standardized measures of writing. The negative correlations between number of sentences and average sentence length reflect that children who tend to write more sentences formulate short sentence. All other significant correlations were positive.

[2] *W* scores were computed for the Woodcock Johnson measures. They were preferred over raw scores when they were available for both indicators of a latent factor because they are more sensitive to increase in ability (Woodcock et al., 2001).

Table 1. Descriptive Statistics

Variable Name	N	Raw Scores M	SD	Min	Max	Z-Scores M	SD	Min	Max
Year 1									
TOWRE NW	315	14.03	8.67	0.00	34.00	0.00	1.00	-1.62	2.30
WJ Word Attack	316	13.76	6.42	3.00	29.00	0.00	1.00	-1.67	2.37
TOSREC Form A	316	18.52	10.88	0.00	42.00	0.00	1.00	-1.70	2.16
TOSREC Form D	315	17.48	10.37	0.00	44.00	0.00	1.00	-1.69	2.56
WJ Passage Comp.	315	463.53+	17.53	415.00	503.00	0.00	1.00	-2.77	2.25
WRMT P. Comp.	316	468.89+	15.46	427.00	500.00	0.00	1.00	-2.71	2.01
WRAT Spelling	316	21.53	3.45	11.00	30.00	0.00	1.00	-3.06	2.46
WIAT Spelling	316	17.30	5.02	4.00	31.00	0.00	1.00	-2.65	2.73
WIAT Sentences	316	0.95	1.12	0.00	5.00	0.00	1.00	-1.64	2.44
Words per Sentence	316	8.23	3.43	1.11	15.88	0.00	1.00	-.85	3.62
No. of Sentences	316	5.84	2.84	1.00	13.00	0.00	1.00	-1.70	2.52
No. of Words	316	44.53	20.18	9.00	97.00	0.00	1.00	-1.76	2.60
Year 2									
TOWRE NW	270	25.75	11.35	1.00	53.00	1.35	1.31	-1.50	4.49
WJ Word Attack	270	19.20	6.11	4.00	31.00	0.85	0.95	-1.52	2.68
TOSREC Form A	270	25.59	9.41	0.00	50.00	0.65	0.87	-1.70	2.89
TOSREC Form D	270	26.00	9.38	0.00	50.00	0.82	0.90	-1.69	3.14
WJ Passage Comp.	270	483.81+	13.73	445.00	515.00	1.16	0.78	-1.06	2.94
WRMT P. Comp.	270	484.76+	11.38	429.00	516.00	1.03	0.74	-2.58	3.05
WRAT Spelling	270	24.75	3.91	15.00	36.00	0.93	1.14	-1.89	4.20
WIAT Spelling	269	23.61	5.52	11.00	38.00	1.26	1.10	-1.26	4.13
WIAT Sentences	270	2.30	1.74	0.00	6.00	1.36	1.14	-.83	4.31
Words per Sentence	270	9.62	3.70	3.00	18.39	0.13	1.17	-1.70	2.70
No. of Sentences	270	6.20	3.31	1.00	13.50	1.21	1.56	-.85	4.52
No. of Words	270	54.93	26.60	9.00	124.50	0.52	1.32	-1.76	3.96
Year 3									
TOWRE NW	260	30.34	10.31	7.00	55.00	1.82	1.27	-.93	4.49
WJ Word Attack	260	21.88	5.65	5.00	31.00	1.26	0.88	-1.36	2.68
TOSREC Form A	260	28.57	8.37	7.00	51.00	0.92	0.77	-1.06	2.99
TOSREC Form D	260	27.96	9.08	0.00	54.00	1.01	0.88	-1.69	3.52
WJ Passage Comp.	260	494.20+	12.17	461.00	521.00	1.75	0.69	-.14	3.28
WRMT P. Comp.	260	494.27+	11.68	430.00	522.00	1.64	0.76	-2.52	3.43
WRAT Spelling	260	28.76	3.73	21.00	43.00	2.10	1.08	-.15	6.23
WIAT Spelling	259	28.30	5.75	16.00	44.00	2.19	1.15	-.26	5.32

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

Table 1. Descriptive Statistics (*continued*)

		Raw Scores				Z-Scores			
Variable Name	N	M	SD	Min	Max	M	SD	Min	Max
WIAT Sentences	260	3.81	2.21	0.00	10.00	0.41	0.76	-1.29	2.21
Words per Sentence	260	10.43	3.26	3.54	18.16	1.28	1.46	-1.70	5.34
No. of Sentences	260	9.47	4.14	1.00	21.00	2.56	1.97	-.85	8.09
No. of Words	260	93.15	36.19	17.00	191.50	2.41	1.79	-1.36	7.28

Table 2. Correlations among variables for the first grade sample (N= 316)

	1	2	3	4	5	6	7	8	9	10	11	12
1. TOWRE	---											
2. WJ WA	0.84	---										
3. TOSREC A	0.79	0.77	---									
4. TOSREC D	0.78	0.75	0.92	---								
5. WJ PC	0.74	0.75	0.84	0.84	---							
6. WRMT PC	0.77	0.78	0.85	0.85	0.87	---						
7. WRAT Sp	0.72	0.76	0.74	0.74	0.71	0.74	---					
8. WIAT Sp	0.70	0.78	0.77	0.79	0.76	0.79	0.80	---				
9. WIAT CS	0.26	0.29	0.33	0.32	0.35	0.36	0.29	0.31	---			
10. ASL	0.24	0.27	0.25	0.25	0.26	0.27	0.24	0.26	0.19	---		
11. NS	$.087^{ns}$	$.082^{ns}$	0.13	0.13	$.073^{ns}$	$.080^{ns}$	0.13	$.092^{ns}$	$-.043^{ns}$	-0.50	---	
12. NW	0.35	0.35	0.41	0.41	0.33	0.36	0.38	0.36	0.12	0.31	0.62	---

Note. TOWRE = Test of Silent Reading Efficiency: Phonetic Decoding Efficiency, WJ WA = Woodcock Johnson Word Attack, TOSREC A = Test of Silent Reading Efficiency and Comprehension Form A, TOSREC D = Test of Silent Reading Efficiency and Comprehension Form D, WJ

^{ns} Correlation is not significant. Correlations greater than .10 are significant at the 0.05 level. Correlations greater than .18 are significant at the 0.001 level.

Table 3. Correlations among variables for the second grade sample (N= 270)

	1	2	3	4	5	6	7	8	9	10	11	12
1. TOWRE	---											
2. WJ WA	0.81	---										
3. TOSREC A	0.68	0.65	---									
4. TOSREC D	0.69	0.65	0.85	---								
5. WJ PC	0.60	0.64	0.75	0.74	---							
6. WRMT PC	0.52	0.56	0.66	0.67	0.73	---						
7. WRAT Sp	0.70	0.74	0.63	0.64	0.62	0.53	---					
8. WIAT Sp	0.64	0.71	0.68	0.67	0.67	0.62	0.83	---				
9. WIAT CS	0.26	0.34	0.35	0.35	0.37	0.35	0.38	0.41	---			
10. ASL	$.097^{ns}$	0.13	$.058^{ns}$	$.040^{ns}$	$.073^{ns}$	$.056^{ns}$	$.095^{ns}$	0.16	0.17	---		
11. NS	$.029^{ns}$	$-.032^{ns}$	0.14	0.15	$.097^{ns}$	$.036^{ns}$	$.029^{ns}$	$-.025^{ns}$	$-.038^{ns}$	-0.49	---	
12. NW	0.12	$.092^{ns}$	0.22	0.22	0.18	$.090^{ns}$	0.12	0.13	$.086^{ns}$	$.108^{ns}$	0.76	---

Note. TOWRE = Test of Silent Reading Efficiency: Phonetic Decoding Efficiency, WJ WA = Woodcock Johnson Word Attack, TOSREC A = Test of Silent Reading Efficiency and Comprehension Form A, TOSREC D = Test of Silent Reading Efficiency and Comprehension Form D, WJ

^{ns} Correlation is not significant. Correlations greater than .10 are significant at the 0.05 level. Correlations greater than .16 are significant at the 0.001 level.

Table 4. Correlations among variables for the third grade sample (N= 260)

	1	2	3	4	5	6	7	8	9	10	11	12
1. TOWRE	---											
2. WJ WA	0.78	---										
3. TOSREC A	0.68	0.60	---									
4. TOSREC D	0.66	0.55	0.82	---								
5. WJ PC	0.62	0.69	0.74	0.66	---							
6. WRMT PC	0.55	0.60	0.67	0.59	0.74	---						
7. WRAT Sp	0.68	0.71	0.65	0.58	0.67	0.60	---					
8. WIAT Sp	0.67	0.69	0.66	0.63	0.66	0.61	0.83	---				
9. WIAT CS	0.47	0.49	0.55	0.53	0.59	0.56	0.46	0.55	---			
10. ASL	0.00	.093^{ns}	.023^{ns}	-.012^{ns}	-.014^{ns}	-.005^{ns}	.092^{ns}	0.14	.038^{ns}	---		
11. NS	0.15	0.13	0.24	0.27	0.28	0.19	0.15	0.18	0.17	-0.48	---	
12. NW	0.21	0.25	0.34	0.34	0.36	0.24	0.29	0.35	0.29	0.14	0.76	---

Note. TOWRE = Test of Silent Reading Efficiency: Phonetic Decoding Efficiency, WJ WA = Woodcock Johnson Word Attack, TOSREC A = Test of Silent Reading Efficiency and Comprehension Form A, TOSREC D = Test of Silent Reading Efficiency and Comprehension Form D, WJ

^{ns} Correlation is not significant. Correlations greater than .10 are significant at the 0.05 level. Correlations greater than .16 are significant at the 0.001 level.

Selection Coh-Metrix Variables

Descriptive Statistics

Table 5 contains descriptive statistics for all of the Coh-Metrix variables. As noted in the table most compositional fluency measures increased per year: number of sentences (READNS), number of words (READNW), average paragraph length (READAPL), and average sentence length (READASL). In addition, the following syntactic, semantic and frequency indices also increased per year: number of words before the main verb (SYNLE), anaphor references between adjacent sentences (CREFP1u), log of the frequency of content words[3] (FRQCLacw) and low frequency score of content words[4] (FRQCRmcs). A number of variables decreased per year indicating a low score was related to more elaborate writing: positive causal connectives (CONCSpi), positive logical connectives (CONLGpi), noun phrases (DENSNP), personal pronouns (DENPRPi), content word overlap between adjacent sentences (CREFC1u), conceptual similarity in adjacent sentences (LSAassa), verb hypernyms (HYVERBaw), lowest concreteness across sentences (WORDCmcs) and causal words (CAUSVP). Of the remaining variables, the means of 13 variables increased in year 2 and decreased in year 3: all connectives (CONi), additive connectives (CONADpi), logical operators (DENLOGi), higher level constituents[5] per word (SYNHw), stem overlap between adjacent sentences (CREFS1u), stem overlap between all sentences (CREFSau), frequency of content words (FRQCRacw), log of the lowest frequency score (FRQCLmcs), average number of syllables per word (READASW), Flesch-Kincaid Grade Level readability (READFKGL), ratio of causal verbs to causal particles[6] (CAUSC), tense and aspect ratio (TEMPta), and location and motion ratio (SPATC). The means of 12 variables decreased in year 2 and increased in year 3: Flesch Reading Ease Score (READFRE), ratio of pronouns to noun phrases (DENSNPR2), modifiers per noun phrase (SYNNP), syntax similarity between adjacent sentences (STRUTa), syntax similarity between all sentences (STRUTt), syntax similarity between sentences within paragraphs (STRUTp), type-token ratio (TYPTOKc), argument overlap between adjacent sentences (CFEFA1u) , argument overlap for all sentences (CREFAau), anaphors (CREFPau), noun hypernym (HYNOUNaw), concreteness

---

[3] Taking the log of the frequencies of content words rather than the raw scores is compatible with reading time.

[4] The low frequency score of content words takes into account the rarest word in sentences, and computes a mean value for each passage. A high score on this measure indicates the use of infrequent content words.

[5] Higher level constituents represent a measure of syntactic density of sentences.

[6] Causal particles are conjunctions, transitional adverbs, and other forms of connectives, such as *since*, *so that*, *because*, and *consequently* that are used to indicate causal relationship between clauses.

(WORDCacw), and intentional words (INTEi). The variability in children's writing was expected because the writing task required expository writing in the first two years and narrative writing in the final year, and because Coh-Metrix is a tool used for analyzing college level writing.

Preliminary Analyses

Preliminary analyses were conducted including multiple regressions, correlations and canonical correlations between Coh-Metrix variables and the standardized measures of reading. Canonical correlations optimize the relation between two sets of variables (e.g., reading and writing) by creating canonical variates, or linear combinations of each set of variables. Each canonical variate is then correlated with the observed variables of the second set (Tabachnick & Fidell, 1996). Subsequent correlations are computed based on the residuals of the first correlation. The results of the canonical correlation showed that the Flesch readability index significantly predicted the reading variate during all years. Thus, the Flesch readability index was broadly related to all of the reading variables including decoding words and non-words, reading sentences and several reading comprehension tasks. The formula for the Flesch readability index is based on two other Coh-Metrix variables: average sentence length and average number of syllables per word.

In addition to readability indices, compositional fluency, syntactic, semantic, frequency and situation model indices were entered as indicators of text or sentence level factors in several re-specifications of the measurement model part of the latent change score models. Following a two step procedure for SEM modeling outlined by Kline (2011), confirmatory factor analyses were conducted to evaluate measurement models at the word, sentence and text levels. Partial measurement invariance was established as a compromise between full measurement invariance and complete lack of invariance. This allowed the use of a scale in which there were some differences in measurement between time points, while still considering the overall comparison to be meaningful. Furthermore, the latent factors were scaled to represent deviations from Year 1 latent factors.

The writing factors at the sentence and text levels consisted of different pairs of Coh-Metrix variables. Some variable pairs were excluded because they provide overlapping information (e.g., positive causal connectives and positive logical connectives both count words such as *so* and *because*) or because they were linear combinations of other variables (e.g., Flesch

Reading Ease score includes average sentence length and average syllables per word). Furthermore, some variables (such as 'words before main verb') were excluded because they were conceptually overlapping at the sentence and text levels (because Coh-Metrix calculates an average value of 'words before main verb' per passage (as opposed to values for sentences nested within passages)). Other variable pairs were excluded because they were conceptually too different.

The results of the preliminary Coh-Metrix analyses showed that in comparison to syntactic, semantic, readability, frequency and situation model indices, compositional fluency indices were more related to reading comprehension, and were better indicators of a text level factor. Thus, the variables that were selected (number of words and number of sentences) were determined to be related yet distinct indicators of compositional fluency. 'Average words per sentence' (or average sentence length) was used as an indicator of a sentence level factor along with the WIAT-II Combining Sentences task.

Table 5. Descriptive Statistics for all Coh-Metrix variables (*z* scores)

	Year 2 (N = 270)		**Year 3 (N = 260)**	
	M	SD	M	SD
Syntactic Indices				
CONADpi	0.24	0.94	0.15	0.68
CONCSpi	-0.09	0.72	-0.63	0.50
CONi	0.19	0.89	0.07	0.73
CONLGpi	-0.10	0.76	-0.34	0.62
DENSNP	-0.09	0.97	-0.16	0.69
DENSPR2	-0.29	0.80	-0.24	0.64
DENLOGi	0.04	0.88	0.02	0.67
DENPRPi	-0.09	0.81	-0.29	0.54
STRUTa	-0.40	0.77	-0.31	0.61
STRUTt	-0.34	0.82	-0.29	0.62
STRUTp	-0.34	0.82	-0.28	0.64
TYPTOKc	-0.23	0.94	-0.11	0.61
SYNNP	-0.11	1.02	0.20	0.71
SYNHw	0.18	0.93	-0.15	0.66
SYNLE	0.29	1.05	0.94	1.17
Semantic Indices				
CREFA1u	-0.10	0.98	-0.04	0.82
CREFS1u	0.09	0.99	-0.29	0.57
CREFP1u	0.01	0.99	0.19	0.74
CREFAau	-0.08	0.91	-0.01	0.75
CREFSau	0.02	0.85	-0.39	0.37
CREFPau	-0.01	0.98	0.05	0.80
CREFC1u	-0.23	0.87	-0.41	0.63
WORDCacw	-1.19	0.78	-0.41	0.61
WORDCmcs	-0.39	0.78	-0.87	0.70
LSAassa	-0.05	0.76	-0.51	0.49
HYNOUNaw	-0.99	0.59	-0.37	0.61
HYVERBaw	-0.10	0.85	-0.27	0.73
Compositional Fluency				
READNS	0.13	1.17	1.28	1.46
READNW	0.52	1.32	2.41	1.79
READAPL	0.13	1.17	0.90	1.49
READASL	0.40	1.08	0.64	0.95
READASW	1.36	1.14	0.41	0.76
Frequency Indices				
FRQCRacw	0.02	0.78	-0.22	0.57
FRQCLacw	0.69	0.75	0.71	0.75

Table 5. Descriptive Statistics for all Coh-Metrix variables (*z* scores) (*continued*)

	Year 2 (N = 270)		**Year 3 (N = 260)**	
	M	SD	M	SD
FRQCRmcs	0.09	1.11	0.66	1.49
FRQCLmcs	0.27	1.24	0.22	0.78
Readability Indices				
READFRE	-2.82	3.07	-0.84	1.64
READFKGL	1.10	1.34	0.64	0.98
Situation Model Indices				
CAUSVP	-0.28	0.91	-0.75	0.57
CAUSC	0.08	0.89	-0.36	0.71
INTEi	-0.47	0.69	0.48	0.84
TEMPta	0.35	0.78	0.09	0.81
SPATC	0.86	0.84	0.74	0.72

Note. CONADpi = Incidence of positive additive connectives, CONCSpi = Incidence of positive causal connectives, CONi = Incidence of all connectives, CONLGpi = Incidence of positive logical connectives, DENSNP = Noun Phrase Incidence Score (per thousand words), DENSPR2 = Ratio of pronouns to noun phrases, DENLOGi = Logical operator incidence score, DENPRPi = Personal pronoun incidence score, STRUTa = Sentence syntax similarity, for adjacent sentences, STRUTt = Sentence syntax similarity, all, across paragraphs, STRUTp = Sentence syntax similarity, sentence all, within paragraphs, TYPTOKc = Type-token ratio for content words, SYNNP = Mean number of modifiers per noun-phrase, SYNHw = Mean number of higher level constituents per word, SYNLE = Mean number of words before the main verb of main clause in sentences, CREFA1u = Argument Overlap for adjacent sentences , CREFS1u = Stem Overlap for adjacent sentences, CREFP1u = Anaphor reference for adjacent sentences, CREFAau = Argument Overlap for all sentences, CREFSau = Stem Overlap for all sentences, CREFPau = Anaphor reference for all sentences, CREFC1u = Proportion of content words that overlap between for adjacent sentences sentences, WORDCacw = Concreteness for content words, WORDCmcs = Concreteness, minimum in sentence for content words, LSAassa = Latent Semantic Analysis (conceptual similarity) for adjacent sentences, HYNOUNaw = Mean hypernym values of nouns, HYVERBaw = Mean hypernym values of verbs, READNP = Number of Paragraphs, READNS = Number of Sentences, READNW = Number of Words, READAPL = Average Sentences per Paragraph, READASL = Average Words per Sentence, READASW = Average Syllables per Word, FRQCRacw = frequency of content words, FRQCLacw = logarithm of the frequency of content words, FRQCRmcs = Celex, lowest frequency score for content words, FRQCLmcs = logarithm of lowest frequency score for content words, READFRE = Flesch Reading Ease Score, READFKGL = Flesch-Kincaid Grade Level, CAUSVP = Incidence of causal verbs, links, and particles, CAUSC = Ratio of causal particles to causal verbs, INTEi = Incidence of intentional actions, events, and particles, TEMPta = Mean of tense and aspect repetition scores, SPATC = Mean of location and motion ratio scores.

Latent Change Score Analysis

A series of latent change score models were conducted using M-plus (Muthen & Muthen, 1998-2011).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, the causal parameters from reading to writing were estimated whereas the causal parameters from writing to reading were fixed to zero. In the writing-to-reading model the causal parameters from writing to reading were estimated and the causal parameter from reading to writing were fixed to zero. In the bidirectional model all causal parameters were estimated. Attempts to improve model fit included correlating the disturbances of observed variables across time points. Furthermore, intercepts were constrained to be equal for three indicators at the word level and four indicators at the sentence and text levels. As the unidirectional models are nested within the bidirectional model, chi-square difference test was employed to statistically compare models. The base model always yields the best fit as it has no 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 models are presented in **Table 6**. All models were identified, indicating that a unique set of parameter estimates were obtained. A non-significant $\chi 2$ value indicates the overall test of model fit is acceptable. Other fit indices were also used to evaluate model fit and compare nested and non-nested models. Values less than 2.0 are desirable for the $\chi 2/df$ test (Kline, 2011). Following the suggestions of Hu and Bentler (1999), a cutoff value of .95 was used as an indicator of good fit for the Tucker-Lewis Index (TLI) and Comparative Fit Index (CFI) indices. A value of less than .07 was used for the RMSEA index (Steiger, 2007) and less than .05 was used for the Standardized Root Mean Square Residual (SRMR) (Byrne, 1998). As observed in **Table 6**, all the models fit well across levels of language. The reading-to-writing models were the most parsimonious, as observed by the low $\chi 2/df$.

Comparison of the unidirectional and bidirectional models with a $\chi 2$ difference test revealed that all reading-to-writing models and the sentence level writing-to-reading model were not significantly different from the bidirectional model (**Table 6**). Because the reading-to-writing models were not significantly different from the bidirectional models and because they were the most parsimonious, they were considered the most appropriate models at all levels of language, and only the results for these models were interpreted in this study.

Table 6. Model Fit Indices

Model	$\chi 2$	df	$\chi 2$/df	CFI	TLI	RMSEA	SRMR	$\chi 2$ difference (vs. bidirectional)
Word Level								
Bidirectional	80.73**	35	2.31	0.99	0.98	0.06	0.07	
R-to-W	83.58**	38	2.20	0.99	0.98	0.06	0.07	2.85
W-to-R	114.78**	38	3.02	0.98	0.97	0.08	0.08	34.05**
Sentence Level								
Bidirectional	70.85**	37	1.91	0.98	0.97	0.07	0.08	
R-to-W	76.45**	40	1.91	0.98	0.97	0.05	0.08	5.60
W-to-R	101.66**	40	2.54	0.97	0.95	0.07	0.09	30.81**
Text Level								
Bidirectional	56.86*	37	1.54	0.99	0.98	0.04	0.04	
R-to-W	57.89*	40	1.45	0.99	0.99	0.03	0.04	1.03
W-to-R	70.35**	40	1.76	0.99	0.98	0.05	0.07	13.49*

* $p < 0.05$. ** $p < 0.01$.

Results for the Reading-to-Writing Models

Evaluations of Mean Structures

Word Level. The estimated parameters of the change scores are summarized in **Table 7**. The mean of the first change in decoding ($\mu_{\Delta r1}$ = 1.24) indicated an increase in decoding between years 1-2. The mean of the second change in decoding was smaller in magnitude but also indicated an increase in decoding ($\mu_{\Delta r2}$ = .52). The means of the first change in spelling ($\mu_{\Delta w1}$ = 1.21) and second change in spelling ($\mu_{\Delta w2}$ = .94) indicated an increase in spelling across years. The variances of changes in decoding ($\sigma^2_{\Delta r1}$ = .215, and $\sigma^2_{\Delta r2}$ = .041) as well as spelling ($\sigma^2_{\Delta w1}$ = .088 and $\sigma^2_{\Delta w2}$ = .03) indicated there was a small amount of interindividual variability in change patterns. The means of the change scores indicated more growth occurred between years 1-2 for both spelling and decoding. Overall, the model significantly accounted for small amount of variability in the first changes in decoding (14%). Conversely, the model accounted for a large amount of variability in the second change in decoding (70%) and second change in spelling (93%).

Sentence Level. The mean of the first and second changes in sentence reading ($\mu_{\Delta r1}$ = .66; and $\mu_{\Delta r2}$ = .26) indicated an increase in sentence reading ability across years. The means of the changes in sentence writing ($\mu_{\Delta w1}$ = 1.18 and $\mu_{\Delta w2}$ =1.36) were larger in magnitude and indicated an increase in sentence writing across years. The variances of the changes in reading ($\sigma^2_{\Delta r1}$ = 17; and $\sigma^2_{\Delta r2}$ = .03) indicated small interindividual differences in growth patterns for sentence reading. The variance of the initial change in sentence writing ($\sigma^2_{\Delta w1}$ = 1.28) indicated a large amount of interindividual variability in sentence writing. The variability of the second change in writing ($\sigma^2_{\Delta w2}$ = .63) indicated small interindividual differences in growth patterns for sentence writing. The means of the change scores indicated that children grew more on sentence writing than sentence reading. Overall, the model significantly accounted for 36% of the variance in first change in sentence reading, 53% of the variance in the first change in sentence writing and 94% of the variance in the second change in sentence writing.

Text Level. The mean of the first change in reading comprehension ($\mu_{\Delta r1}$ = 1.13,) indicated a large increase in reading ability between years 1-2. The mean of the second change in reading comprehension ($\mu_{\Delta r2}$ = .59) indicated a small increase in reading comprehension between years 2-3. Conversely, the mean of the fist change in text writing ($\mu_{\Delta w1}$ = .51) indicated a smaller increase in writing ability between years 1 -2. The mean of the second change in text writing ($\mu_{\Delta w2}$ = 1.86) indicated a larger increase between years 2 -3. The variances of the changes in text reading ($\sigma^2_{\Delta r1}$ = . .15 and $\sigma^2_{\Delta r2}$ = .01) indicated small interindividual differences in growth patterns for text reading. The variances for the changes in text writing ($\sigma^2_{\Delta w1}$ = 6.51and $\sigma^2_{\Delta w2}$ = 8.17) indicated large interindividual variability in all changes in text writing. The means of the change scores indicated that children grew more on text level reading between years 1-2. Conversely, children grew more on writing between years 2-3 than years 1-2. Overall, the model significantly accounted for 55% of the variance in the first change in reading comprehension, and 45% of the variance in the second change in text writing.

Table 7. Estimated Means, Standard Deviations, Minimum, Maximum and R-square values of Change Scores for the Reading-to-Writing models

	Δ1					Δ2				
	M	SD	Min	Max	R^2	M	SD	Min	Max	R^2
Reading										
Word	1.24	0.46	-0.04	2.40	.14*	0.52	0.20	-0.01	1.08	.70*
Sentence	0.66	0.41	-0.55	2.27	.36**	0.26	0.18	-0.33	0.89	.15
Text	1.13	0.39	0.08	2.39	.55**	0.59	0.12	0.21	1.09	.22
Writing										
Word	1.21	0.30	0.50	2.44	.11	0.94	0.17	0.51	1.50	.93*
Sentence	1.18	1.13	-3.44	4.09	.53*	1.36	0.79	-0.44	3.65	.94*
Text	0.51	2.55	-5.94	8.90	.05	1.86	2.86	-8.03	8.53	.45**

* $p < 0.05$, ** $p < 0.001$

Evaluations of the Structural Model

Word Level. Figure 3 contains the unstandardized regression coefficients for the word level reading-to-writing model. Factor correlation at year 1 showed that decoding was highly correlated with spelling (r = .93, $p < 0.001$). Findings showed that change in decoding between years 1-2 was predicted by decoding at year 1 ($\beta = .21$, $p < .001$), indicating that children who scored high on decoding changed between years 1-2. The change in decoding between years 2-3 was predicted by decoding at year 2 ($\beta = -.24$, $p < 0.001$) as well as the change in decoding between years 1-2 ($\alpha = .47$, $p < 0.05$). These findings indicate that children who scored low on year 2 decoding changed between years 2-3. In addition, children who changed in decoding between years 1-2 also changed in decoding between years 2-3.

In terms of bivariate effects, change in spelling between years 1-2 was predicted by year 1 decoding ($\gamma = .37$, $p < 0.05$), after controlling for year 1 spelling ($\beta = -.31$, $p < 0.05$). These findings suggest that children who scored low on year 1 spelling changed between yeas 1-2. In addition, children who scored high on decoding changed in spelling between years 1-2. Change in spelling between years 2-3 was predicted by change in decoding between years 1-2 ($\theta = .66$, $p < 0.001$) and year 2 decoding ($\gamma = -.26$, $p < 0.05$), after controlling for change in spelling between years 1-2 ($\alpha = -.32$, $p = 0.06$) and year 2 spelling ($\beta = .16$, $p = .26$). These findings illustrate the dynamic relations between decoding and spelling. Specifically, children who had changed on decoding between years 1-2 also changed on spelling between years 2-3. Furthermore, and children who changed less on year 2 decoding changed in spelling between years 2-3. Finally, there was a possible indirect effect on the change in spelling between years 2-3. The first change in decoding could possibly mediate the effect of year 1 decoding on the change in spelling between years 2-3. However, this effect is not reported because the model did not include a direct effect pathway between year 1 decoding and the second change in spelling.

Sentence Level. Figure 4 contains the unstandardized regression coefficients for the sentence level reading-to-writing model. The correlation of .37 ($p < .001$) indicates a moderate relation between the initial reading and writing factors. Findings showed that change in reading between years 1-2 was predicted by reading at year 1 ($\beta = -.30$, $p < .001$). Specifically, children who scored low on reading at year 1 changed between years 1-2. Change in reading between years 2-3 was predicted by reading at year 2 ($\beta = -.12$, $p < 0.05$) and was not predicted by the change in reading between years 1-2 (α = .08, p = 0.27), indicating that children who scored low on reading at year 2 changed between years 2-3.

In terms of bivariate effects, change in sentence writing between years 1-2 was predicted by year 1 reading (γ = .65, $p < 0.001$), and year 1 writing (β = -.98, $p < 0.001$). These findings suggest that children who scored high on year 1 sentence reading changed on sentence writing between years 1-2. Furthermore, children who scored low on sentence writing at year 1 changed on sentence writing between years 1-2. Change in writing between years 2-3 was predicted by reading at year 2 (γ = 1.43, $p < .01$) and year 2 writing (β = -1.05, $p < 0.05$), after controlling for change in writing between years 1-2 (α = .001, p = 0.99) as well as change in reading between years 1-2 (θ = -.36, p = 0.36). Similar to the change in sentence writing between years 1-2, these findings indicate that children who scored high on reading during year 2 changed on writing between years 2-3, and children who scored low on writing at year 2 changed in writing between years 2-3.

Text Level. Figure 5 contains the unstandardized regression coefficients for the text level reading-to-writing model. The correlation of .37 ($p < .001$) indicated a moderate relation between the initial reading and writing factors at the text level. Findings showed that change in reading between years 1-2 was predicted by reading at year 1 (β = -.36, $p < .001$). As expected, children who scored low on reading at year 1 improved between years 1-2. Similarly, the change in reading between years 2-3 was predicted by reading at year 2 (β = -.16, $p < 0.05$), indicating that children who scored low on reading at year 2 improved between years 2-3. The change in reading between years 2-3 was not predicted by the change in reading between years 1-2 (α = -.11, p = 0.27).

Change in writing between years 1-2 was not predicted by year 1 reading (γ = .004, p = 0.68), but was predicted by year 1 writing (β = -.42, $p < .05$). Similar to reading, children who scored low on writing at year 1 changed in text level writing between years 1-2. Change in writing between years 2-3 was predicted by year 2 reading (γ = .58, $p < 0.05$), after controlling for change in writing between years 1-2 (α = -.50, $p < 0.001$) and year 2 writing (β = -.33, $p < 0.05$), after controlling for change in reading between years 1-2 (θ = -.01, p = .98),. These findings suggest that children who scored high on reading during year 2 changed on writing between years 2-3, and children who scored low on writing during year 2 changed on writing between years 2-3.

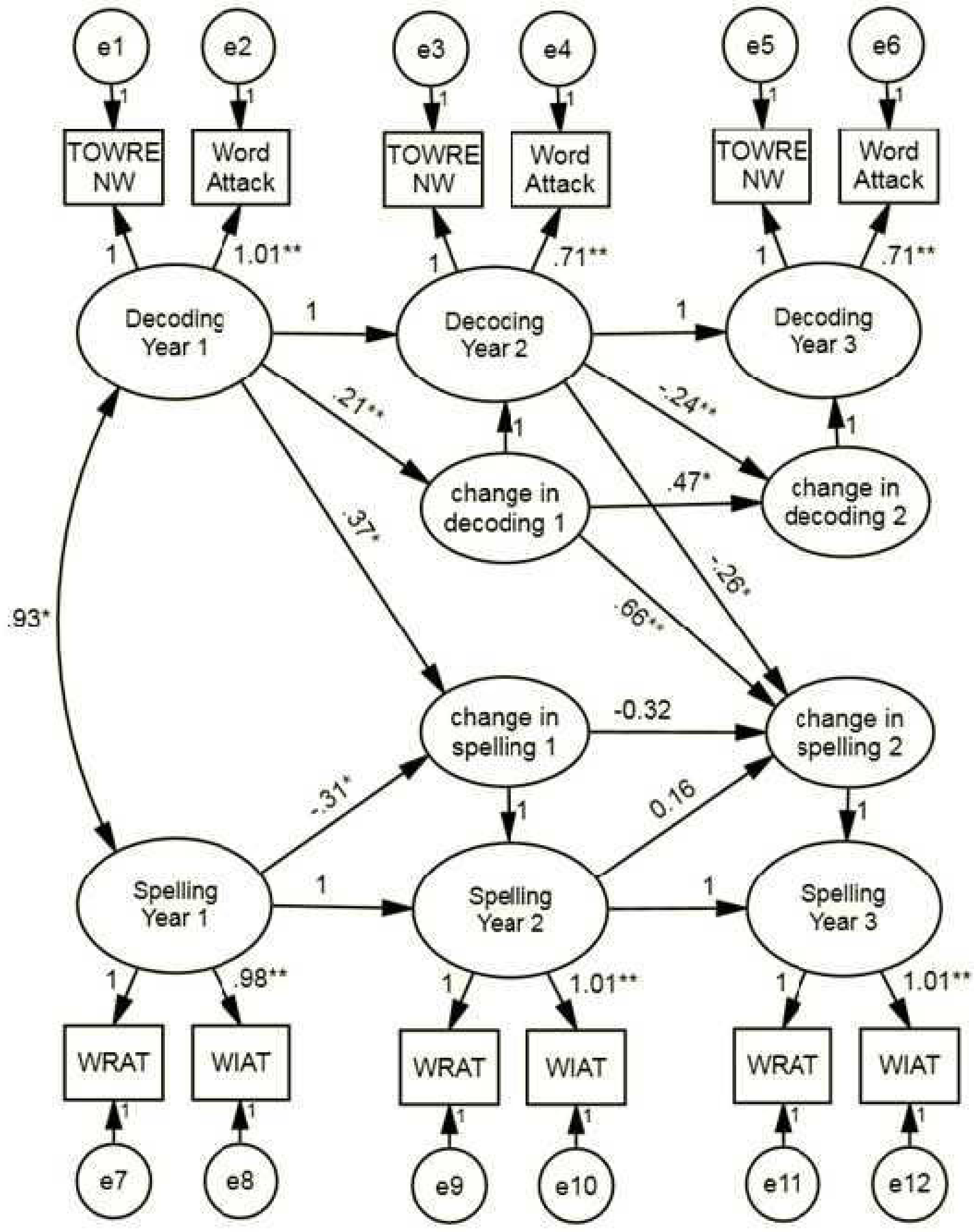

Figure 3. Unstandardized Estimates for the Word Level Reading-to-Writing Model.

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

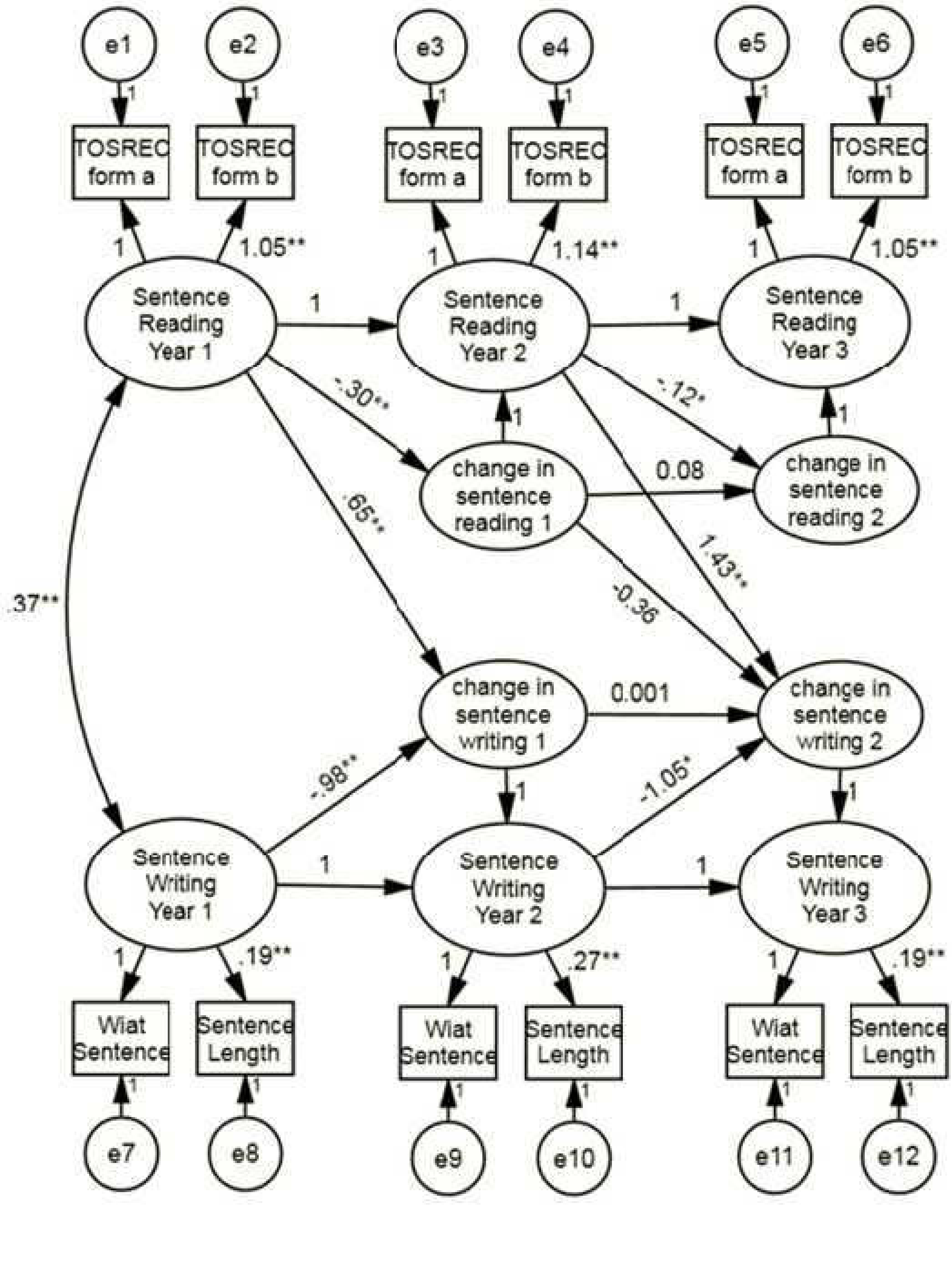

Figure 4. Unstandardized Estimates for the Sentence Level Reading-to-Writing Model.
* $p < 0.05$. ** $p < 0.001$

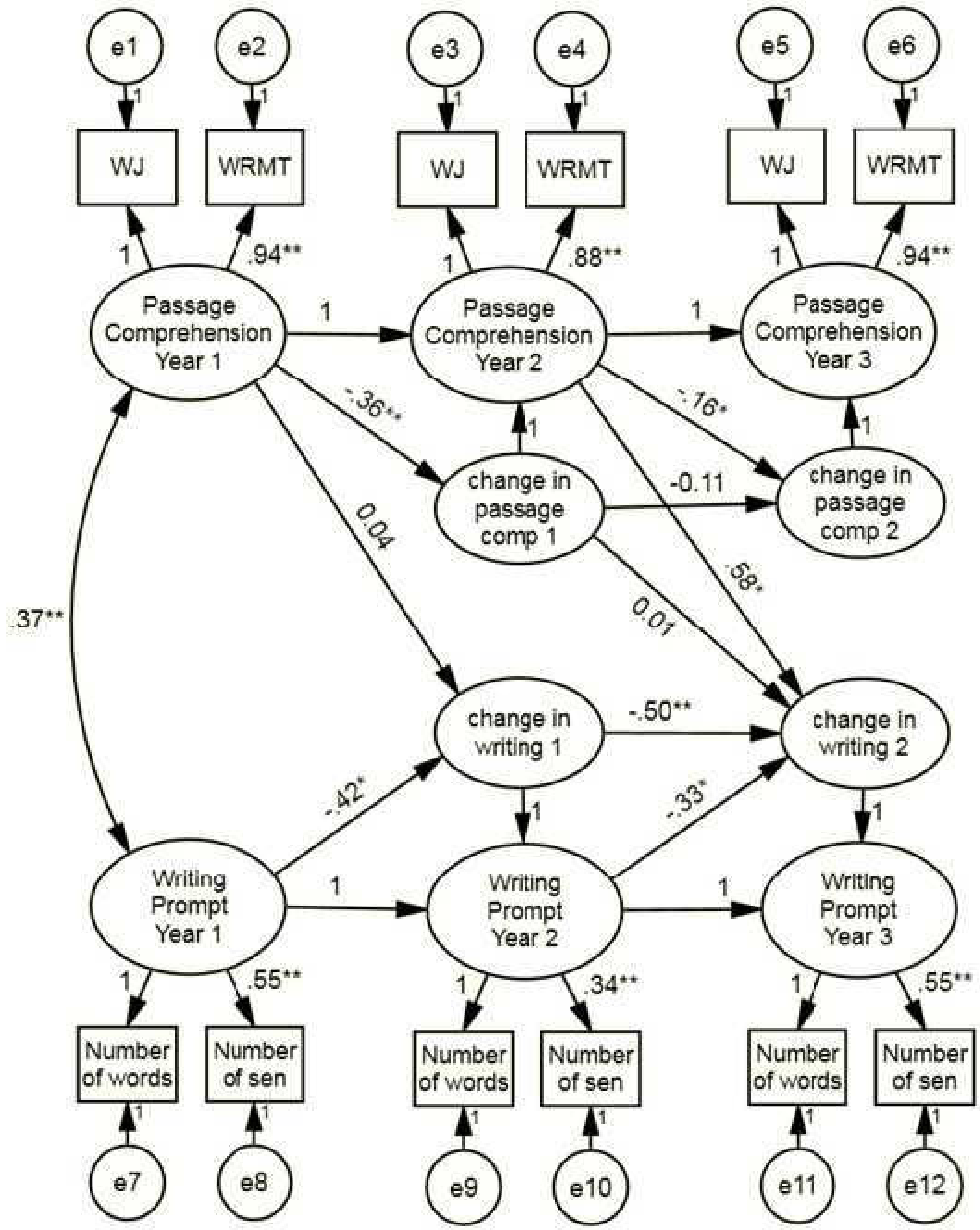

Figure 5. Unstandardized Estimates for the Text Level Reading-to-Writing Model.
* $p < 0.05$. ** $p < 0.001$

DISCUSSION

The purpose of the current study was to utilize latent change score modeling to investigate the unidirectional and bidirectional relations between reading skills (decoding, sentence reading and text comprehension) and writing skills (spelling, writing sentences and writing paragraphs). The use of dynamic modeling allowed us to investigate the nature of the growth in reading and writing, 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.

Although research shows that bidirectional relations between reading and writing exist (Lerkkanen et al., 2004; Shanahan & Lomax, 1986, 1988; Abbott et al., 2010), the present study found that reading-to-writing models were superior to writing-to-reading and bidirectional models. Our findings suggest that reading exerts a relatively larger influence on writing factors than the influence of writing on reading factors. 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).

Although there are several aspects of written language that can be assessed, the present study defined writing as spelling at the word level, sentence construction at the sentence level and compositional fluency at the text or passage level. Thus, it is possible that when assessing different aspects of writing (such as handwriting or qualitative ratings of overall writing quality) different results might hold. Furthermore, our results need to be interpreted in the context of the general education framework. Firstly, reading instruction is more widespread than 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 equally 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 in the learning disability framework are examined. Hence, our results apply to English speaking first through third grade students, from an upper-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 more valid description of longitudinal data than latent growth curve models and cross-lagged structural equation models (Ferrer & McArdle, 2010).

The findings of this study showed that it's possible to disentangle the variability in change from overall achievement status. As expected, the results from the LCS analysis suggested that intraindividual changes in reading and writing were characterized by a rise in scores, although the variances of the change scores suggested interindividual variability in the pattern of changes. Our results indicated that children grew more during earlier stages of development (between grades 1-2) on reading at all levels of language. This may be indicative of developmental processes whereby children learn to read between grades 1 and 2 more so than between grades 2-3. Similarly, more growth occurred for spelling between grades 1-2 than grades 2-3. The opposite pattern was found for sentence and text level writing, with more growth occurring between years 2-3. These findings are consistent with theories of development of writing, whereby development of higher level aspects of writing follow development of word and phrase generation (Fitzgerald & Shanahan, 2006).

The correlations between reading and writing at year 1 were of particular interest. The relation was strong at the word level (r=.93) and moderate at the sentence level (r=.37) and at the text levels (r = .37). 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 word reading and spelling form a common literacy factor at the word level.

Univariate reading-to-reading and writing-to-writing pathways were also examined in this study. Consistent with the lag hypothesis that most children eventually become proficient readers (e.g., Stanovich et al., 1988), low status at the previous year predicted subsequent changes in reading at all levels of language. Hence, as expected, children who were growing were those who had a low status on reading or writing achievement the previous year. As an exception to this pattern, high status in grade 1 decoding predicted initial change in decoding, suggesting the presence of a Mathew effect. Matthew effects suggest that skilled readers learn more and grow more than less skilled readers (Stanovich, 1986), as a function of exposure to and experience with reading. Thus, similar to previous research that showed that phonological processing predicts subsequent decoding (e.g., Sprenger-Charolles et al., 2003; Caravolas et al., 2005) these findings show that mastery on phonological decoding in grade 1 is conducive to improvement on decoding between grades 1 and 2. Our findings also indicated that change in decoding between grades 2-3 was a function of initial change in decoding. This pattern is also characterized by a Matthew effect, whereby children who improved on decoding in grades 1-2 kept improving between grades 2-3. Mathew effects were not found for the changes in sentence and text level reading, suggesting that high scores are not conducive to subsequent growth in sentence and passage level reading.

In terms of writing-to-writing pathways, the developmental lag hypothesis held for all changes (with the exception of the second change in spelling), whereby changes in writing were predicted by their low achievement status at the previous year. Similar to the reading-to-reading pathways, this finding suggested that children who lagged on writing at all levels of language improved between grades 1 and 2, and children who lagged on higher level writing improved between grades 2 and 3. The absence of mastery learning for the change in spelling between grades 2-3 underscores the strong relationship between word level reading and spelling, over and above previous spelling ability. Our results also showed that the second changes spelling and text writing were predicted by low initial change scores. This suggested that children who did not grow between grades 1 and 2 grew between grades 2 and 3. Mathew effects were not found for changes in writing ability.

Next we turn to the multivariate reading-to-writing pathways. Changes in writing were predicted by either achievement status or growth (change) in reading. At the word level, high status on grade 1 decoding predicted change in spelling, suggesting that children who start high on decoding in grade 1 improve on spelling between grades 1-2. In addition, the second change in spelling was predicted by the first change in decoding, suggesting that children who improved on decoding between grades 1-2 improved on spelling between grades 2-3. 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., 2004; 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). Our findings also showed that an improvement in decoding leads to an improvement in spelling. This finding is consistent with spelling interventions that are based on word and pseudo-word recognition (Shanahan, 2006).

At the sentence level, change in sentence construction across years was a function of high status in sentence reading at the previous year. Thus, our findings suggest that the ability to read sentences correctly facilitates writing them correctly. 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 results of the preliminary Coh-Metrix analyses showed that compositional fluency indices were more related to reading comprehension than other indices of writing ability. This is consistent with the findings from other studies (e.g., Lerkkanen et al., 2004; Berninger et al., 2002) that suggest that amount of writing done is a robust measure of writing ability. This could also explain why the model fit was the best for the text level models. At the text level, the change in compositional fluency between grades 2-3 was a function of high status in grade 2 reading. This suggests that, during later periods of development, passage comprehension is predictive of growth in compositional fluency. One possible explanation is that children who read for comprehension are able to write more and therefore convey more ideas than children who are not aware of how written language is used to convey meaning. Overall, our findings suggest that the longitudinal relations among word level variables were the strongest, as an improvement in reading lead to an improvement in subsequent writing. At the sentence level high status in reading lead to an improvement in sentence construction across the years. Finally, at the text level, only high status in grade 2 reading lead to an improvement in compositional fluency between grades 2-3. These findings are consistent with previous research which suggests that lower level connections between literacy skills are stronger than higher level connections (Shanahan, 2006).

Further exploration of a within-levels of language approach would be a productive area for future research (e.g., reading comprehension could predict sentence writing, or word reading could predict compositional fluency). It is also possible that models will be different based on whether children increase in both reading and writing, are discrepant in their reading and writing growth or are poor at both. Thus, future research should examine group differences between children. Furthermore, our models accounted for a low amount of variability in the changes. This suggests the need to add more predictors of reading and writing such as vocabulary, working memory, and language comprehension.

In conclusion, this study extends the literature on developmental relations between reading and writing with four findings. First, we have shown that a reading-to-writing model is more adequate than a bidirectional model, and this relation holds across levels of language. Second, latent change score models are adequate for describing growth in reading and writing, and relations between them. Third, changes in writing are predicted by high status in reading at all levels of language, and this effect was significant at the earlier stage of development at the word and sentence levels, and at later stages of development at the text level. Fourth, improvement in spelling is predicted by an improvement in decoding between grades 1-2. Thus, our results show reading is a leading indicator of writing at word level, and that Mathew effects are characteristic of reading-writing development in that acquisition of writing skills is facilitated for good readers. 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.

APPENDIX A

WRITING PROMPTS

<u>Grade 1 (choosing a classroom pet)</u>

Uncorrected sample

I Wut a Tigre to cile ethre PePle and Tetce and I wut to be King of the Tegres. and I can be srog or I have srips!

Corrected for spelling

I want a tiger to kill other people and teacher and I want to be king of the tigers. And I can be strong or I have stripes!

Corrected for spelling, punctuation, usage, grammar, and syntax

I want a tiger to kill other people and the teacher and I want to be king of the tigers. And I can be strong or have stripes!

<u>Grade 2 (favorite subject)</u>

Uncorrected sample

Reading is my favorite subjetct. Because I like to see the pictues in the book and read them. I like to do math too. Because I am really smart. When my teacher tells me to do some math, I do it really qick.

Corrected for spelling

Reading is my favorite subject. Because I like to see the pictures in the book and read them. I like to do math too. Because I am really smart. When my teacher tells me to do some math, I do it really quick.

Corrected for spelling, punctuation, usage, grammar, and syntax

Reading is my favorite subject because I like to see the pictures in the book and read them. I like to do math too because I am really smart. When my teacher tells me to do some math, I do it really quickly.

Grade 3 (a day off from school)

Uncorrected sample

One day when I had a day off from school, I went to Wild Adventures. A friend came along with us too. When we were outside the gate of Wild Adventures, me and my friend saw a roller coaster. The went up like 60 feet high! Me and my friend raced to gate. We got in found a few rides. We went on a roller coaster, bumper cars, and some other stuff. Then we all went to lunch and then headed into the water parks. There was more rides, so we rode some more rides. After we were done, we went on one more ride and went home.

Corrected for spelling

One day when I had a day off from school, I went to Wild Adventures. A friend came along with us too. When we were outside the gate of Wild Adventures, me and my friend saw a roller coaster. The went up like 60 feet high! Me and my friend raced to gate. We got in found a few rides. We went on a roller coaster, bumper cars, and some other stuff. Then we all went to lunch and then headed into the water parks. There was more rides, so we rode some more rides. After we were done, we went on one more ride and went home.

Corrected for spelling, punctuation, usage, grammar, and syntax

One day when I had a day off from school, I went to Wild Adventures. A friend came along with us too. When we were outside the gate of Wild Adventures, my friend and I saw a roller coaster. They went up like 60 feet high! My friend and I raced to the gate. We got in and found a few rides. We went on a roller coaster, bumper cars, and some other stuff. Then we all went to lunch and then headed into the water parks. There were more rides, so we rode some more rides. After we were done, we went on one more ride and went home.

APPENDIX B

COH METRIX VARIABLES

No.	Measure	Description
1	Adjacent anaphor reference	Anaphor reference, adjacent, unweighted
2	Anaphor reference	Anaphor reference, all distances, unweighted
3	Adjacent argument overlap	Argument Overlap, adjacent, unweighted
4	Argument overlap	Argument Overlap, all distances, unweighted
5	Adjacent stem overlap	Stem Overlap, adjacent, unweighted
6	Stem overlap	Stem Overlap, all distances, unweighted
7	Content word overlap	Proportion of content words that overlap between adjacent sentences
8	LSA sentence adjacent	LSA, Sentence to Sentence, adjacent, mean
9	LSA sentence all	LSA, sentences, all combinations, mean
10	LSA paragraph	LSA, Paragraph to Paragraph, mean
11	Personal pronouns	Personal pronoun incidence score
12	Pronoun ratio	Ratio of pronouns to noun phrases
13	Type-token ratio	Type-token ratio for all content words
14	Causal content	Incidence of causal verbs, links, and particles
15	Causal cohesion	Ratio of causal particles to causal verbs
16	Intentional content	Incidence of intentional actions, events, and particles.
17	Intentional cohesion	Ratio of intentional particles to intentional content
18	Syntactic structure similarity adjacent	Sentence syntax similarity, adjacent
19	Syntactic structure similarity all-1	Sentence syntax similarity, all, across paragraphs
20	Syntactic structure similarity all 2	Sentence syntax similarity, sentence all, within paragraphs
21	Temporal cohesion	Mean of tense and aspect repetition scores
22	Spatial cohesion	Mean of location and motion ratio scores.
23	All connectives	Incidence of all connectives
24	Conditional operators	Number of conditional expressions, incidence score
25	Pos. additive connectives	Incidence of positive additive connectives
26	Pos. temporal connectives	Incidence of positive temporal connectives
27	Pos. causal connectives	Incidence of positive causal connectives
28	Pos. logical connectives	Incidence of positive logical connectives
29	Neg. additive connectives	Incidence of negative additive connectives
30	Neg. temporal connectives	Incidence of negative temporal connectives
31	Neg. causal connectives	Incidence of negative causal connectives
32	Neg.logical connectives	Incidence of negative logical connectives
33	Logic operators	Logical operator incidence score (and + if + or + cond + neg)

34	Raw freq. content words	Celex, raw, mean for content words (0-1,000,000)
35	Log freq. content words	Celex, logarithm, mean for content words (0-6)
36	Min. raw freq. content words	Celex, raw, minimum in sentence for content words (0-1,000,000)
37	Log min. freq. content words	Celex, logarithm, minimum in sentence for content words (0-6)
38	Concreteness content words	Concreteness, mean for content words
39	Min. concreteness content words	Concreteness, minimum in sentence for content words
40	Noun hypernym	Mean hypernym values of nouns
41	Verb hypernym	Mean hypernym values of verbs
42	Negations	Number of negations, incidence score
43	NP incidence	Noun Phrase Incidence Score (per thousand words)
44	Modifiers per NP	Mean number of modifiers per noun-phrase
45	Higher level constituents	Mean number of higher level constituents per word
46	Words before main verb	Mean number of words before the main verb of main clause in sentences
47	No. of words	Number of Words
48	No. of sentences	Number of Sentences
49	No. of paragraphs	Number of Paragraphs
50	Syllables per word	Average Syllables per Word
51	Words per sentence	Average Words per Sentence
52	Sentences per paragraph	Average Sentences per Paragraph
53	Flesch Reading Ease	Flesch Reading Ease Score (0-100)
54	Flesch-Kincaid	Flesch-Kincaid Grade Level (0-12)

APPENDIX C

INSTITUTIONAL REVIEW BOARD APPROVAL

Office of the Vice President For Research
Human Subjects Committee
Tallahassee, Florida 32306-2742
(850) 644-8673, FAX (850) 644-4392

RE-APPROVAL MEMORANDUM

Date: 8/12/2010

To: Richard Wagner [rkwagner@psy.fsu.edu]

Address: 4301
Dept.: PSYCHOLOGY DEPARTMENT

From: Thomas L. Jacobson, Chair

Re: Re-approval of Use of Human subjects in Research
The Florida Longitudinal Study

Your request to continue the research project listed above involving human subjects has been approved by the Human Subjects Committee. If your project has not been completed by 8/10/2011, you are must request renewed approval by the Committee.

If you submitted a proposed consent form with your renewal request, the approved stamped consent form is attached to this re-approval notice. Only the stamped version of the consent form may be used in recruiting of research subjects. You are reminded that any change in protocol for this project must be reviewed and approved by the Committee prior to implementation of the proposed change in the protocol. A protocol change/amendment form is required to be submitted for approval by the Committee. In addition, federal regulations require that the Principal Investigator promptly report in writing, any unanticipated problems or adverse events involving risks to research subjects or others.

By copy of this memorandum, the Chair of your department and/or your major professor are reminded of their responsibility for being informed concerning research projects involving human subjects in their department. They are advised to review the protocols as often as necessary to insure that the project is being conducted in compliance with our institution and with DHHS regulations.

Cc: Janet Kistner, Chair [kistner@psy.fsu.edu]
HSC No. 2010.4749

APPENDIX D

PARENT CONSENT FORM

The Florida State University
Department of Psychology
1107 W Call St.
Tallahassee, Florida 32306-4301
Voice 850/644-2040 Fax 850/644-7739

PLEASE SIGN AND RETURN TO TEACHER

Dear ____________________.

I am a professor in the Department of Psychology at Florida State University. I am conducting a research study to help us better understand the kinds of reading skills that promote reading comprehension. We are asking parents of all children in the first grade at your school to consent to have their child participate in the study.

This is a four-year study. Once every year, your child's participation will involve reading silently, reading out loud, writing, listening to reading passages, and tasks that most children find enjoyable. One of the reading out loud tasks will be recorded. We will also look at the "Write Upon Request" writing assignment that is a part of your child's regular school work. The tasks will be given individually in an open area at your child's school to children whose parents or guardians indicate consent. The total time involved is about 75 minutes.

The participation of your child in this study is voluntary. Your consent may be withdrawn at any time. There will be no penalty and it will not affect your child's grade. Your child can also choose not to participate without penalty.

The results of the study may be published, but your child's name will not be used. Confidentiality of records will be maintained to the full extend allowed by Florida law. Possible benefits of your child's participation include increased understanding of the kinds of reading skills that are needed for successful reading comprehension.

If you have any questions concerning this study or your participation, please contact my research coordinator, Dr. Liz Foster, at foster@psy.fsyu.edu or at 850-645-7428, or me at rkwagner@psy.fsu.edu or at 850-644-1033.

Sincerely,

Richard K. Wagner
Alfred Binet Professor of Psychology

I GIVE permission for my child, ______________________ to participate in the above study.

Parent's Name: ____________________________

Parent's Signature ____________________________

Address ____________________________ Date: ____________

FLORIDA STATE UNIVERSITY
Approved 7/20/07
07.171-R
3/12/08
INSTITUTIONAL REVIEW BOARD

If you have any questions about your child's rights as a participant in this research, of if you feel you have been placed at risk, you can contact the Chair of the Human Subject's Committee, Institutional Review Board, through the Vice President for the Office of Research at (850) 644-8633.

REFERENCES

Abbott, R. D., & Berninger, V. W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary- and intermediate-grade writers. *Journal of Educational Psychology, 85(3)*, *85*, 478-508.

Abbott, R. D., Berninger, V. W. & Fayol (2010). Longitudinal Relationships of Levels of Language in Writing and Between Writing and Reading in Grades 1 to 7. *Journal of Educational Psychology, 102 (3),* 281- 298.

Aarnoutse, Car, van Leeuwe, Jan & Verhoeven, Ludo (2005). Early literacy from a longitudinal perspective. *Educational Research and Evolution, 11* (3), 253-275.

Babayigit, Selma & Stainthrop, Rhona (2011). Modeling the relationships between cognitive-linguistic skills and literacy skills: new insights from a transparent orthography. *Journal of Educational Psychology, 103* (1), 169-189.

Bates, E., & Goodman, J. (1999). The emergence of grammar from the lexicon. In B. MacWhinney (Ed.), *The emergence of language* (pp. 27–80). Mahwah, NJ: Lawrence Erlbaum Associates.

Beers, S.F. & Nagy, W. E. (2011). Writing development in four genres from grade three to seven: syntactic complexity and syntactic generation. *Reading and Writing, 24*, 183-202.

Berninger, V. W., Abbott, R. D., Abbott, S. P., Graham, S., & Richards, T. (2002). Writing and reading: Connections between language by hand and language by eye. *Journal of Learning Disabilities. Special Issue: The language of written language. Vol 35(1)*, *35*(2002), 39-56.

Berninger, V. W., Vaughan, K., Abbott, R., Begay, K., Byrd, K., Curtin, G. (2002b). Teaching spelling and composition alone and together: Implications for the simple view of writing. *Journal of Educational Psychology, 94,* 291-304.

Berninger, V. W., Abbott, R. D., & Alsdorf, B. J. (1997). Lexical- and sentence-level processes in comprehension of written sentences. *Reading and Writing. Vol 9(2)*, *9*, 135-162.

Berninger, V., Abbott, R., Swanson, H. L., Lovitt, D., Trivedi, P., Lin, S., & Amtmann, D. (2010). 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, 41, 179-193.

Byrne, B.M. (1998). *Structural Equation Modeling with LISREL, PRELIS and SIMPLIS: Basic Concepts, Applications and Programming*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Caravolas, M., Hulme, C., & Snowling, M. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. Journal of Memory and Language, 45, 751-774.

Cataldo, S., & Ellis, N. (1988). Interactions in the development of spelling, reading and phonological skills, *Journal of Research in Reading, 11,*86–109.

Dell, (1986). A spreading-activation theory of retrieval sentence production. *Psychological Review, 93* (3), 283-321.

Ehri, L. C. (2005). Development of sight word reading: Phases and findings. In M. J. Snowling & C.Hulme (Eds.), *The science of reading* (pp. 135–155). Oxford: Blackwell.

Ferrer, Emilio & McArdle John J. (2010). Longitudinal Modeling of Developmental Changes in Psychological Research. *Current Directions in Psychological Science, 19 (3),* 149 - 154.

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist. Special Issue: Writing development: The role of cognitive*, *35*(2000), 39-50.

Foorman, B. R., Francis, D. J., Novy, D. M., & Liberman, D. (1991). How letter-sound instruction mediates progress in first-grade reading and spelling. *Journal of Educational Psychology, 83*, 456-469.

Foorman, B. R., Arndt, E. J., & Crawford, E. C. (2011). Important constructs in literacy learning across disciplines. *Topics in Language Disorders, 31* (1), 73-83.

Graesser, Arthur C.; Mcnamara, Danielle S.; Louwerse, Max M. & Cai, Zhiqiang (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers,* 36 (2), 193-202.

Hayes, J. R., & Berninger, V. (2010). Relationships between idea generation and transcription: How act of writing shapes what children write. In C. Brazerman, R. Krut, K. Lunsford, S. McLeod, S. Null, P. Rogers, & A. Stansell (Eds.), *Traditions of Writing Research* (pp. 166-180). New York: Taylor & Frances/Routledge.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6* (1), 1-55.

Jenkins, J. R., Johnson, E., & Hileman, J. (2004). When is reading also writing: Sources of individual differences on the new reading performance assessments. *Scientific Studies of Reading, 8(2)*, 25-151.

Juel, C., Griffith, P., & Gough, P. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. Journal of Educational Psychology, 78, 243-255.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. Journal of Educational Psychology, 80, 437-447.

Kline, R.B. (2011). *Principles and Practice of Structural Equation Modeling* (3rd Edition). New York: The Guilford Press.

Langer, J. A., & Flihan, S. (2000). *Writing and reading relationships: Constructive tasks.* Indrisano, Roselmina; Squire, James R. (eds.) Perspectives on writing: research, theory, and practice. Newark: International Reading Association.

Lerkkanen, M., Rasku-Puttonen, H., Aunola, K., & Nurmi, J. (2004). The Developmental Dynamics of Literacy Skills During the First Grade. *Educational Psychology. Vol 24(6)*, *24*, 793-810.

Little, R.J.A., & Rubin, D.B. (1989). The analysis of social science data with missing values. *Sociological Methods and Research, 18,* 292-326.

Loban, W. D. (1963). *The language of elementary school children*. Urbana, IL: National Council of Teachers of English.

McArdle, John J. (2009). Latent Variable Modeling of Differences and Changes with Longitudinal Data. *Annual Review of Psychology, 60,* 577- 603.

McArdle , John J. & Grimm, Kevin J. (2010). Five Steps in Latent Curve and Latent Change Score Modeling with Longitudinal Data. *Longitudinal Research with Latent Variables*, 245-273.

McNamara, D.S., Louwerse, M.M., Cai, Z., & Graesser, A. (2005). Coh-Metrix version 2.1. Retrieved [10/11/2010], from http//:cohmetrix.memphis.edu.

Mehta, P., Foorman, B., Branum-Martin, L., & Taylor, W. (2005). 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, 9,* 85–116.

Muthén, L., & Muthén, B. (1998-2011). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Nelson, N. W., & Van Meter, A. M. (2007). Measuring written language ability in narrative samples. *Reading & Writing Quarterly, 23*, 287-309.

Puranik, C. S., Wagner, R. K. , Young-Suk, K. & Lopez, D. (2011) Describing Development of Translation from Early to Middle Childhood.

Puranik, C. & Lonigan, C. (2010). From scribbles to scrabble: Preschool children's developing knowledge of written language. *Reading and Writing*, 24, 567-589.

Rode, S. S. (1975). Development of phrase and clause boundary reading in children. *Reading Research Quarterly, 10* (1), 124-142.

Shanahan, T. (1984). Nature of reading-writing relations: An exploratory multivariate analysis. *Journal of Educational Psychology, 76,* 466-477.

Shanahan, T. (2006). *Relations among Oral Language, Reading, and Writing Development.* In MacArthur, Charles A.; Graham, Steve; Fitzgerald, Jill (eds.). Handbook of writing research. (pp. 171-183). New York, NY, US: Guilford Press. xi.

Shanahan, T., & Lomax, R. G. (1986). An analysis and comparison of theoretical models of the reading–writing relationship. *Journal of Educational Psychology. Vol 78(2)*, *78*, 116-123.

Shanahan, T., & Lomax, R. G. (1988). A Developmental Comparison of Three Theoretical Models of the Reading-Writing Relationship. *Research in the Teaching of English*, *22*(2), 196-212.

Sprenger-Charolles, L., Siegel, L. S., Bechennec, D. & Serniclaes, W. (2003). 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, 84*, 194-217.

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences, 42* (5), 893-98.

Stanovich, K. E. (1986). Mathew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21* (4), 360-407.

Stanovich, K. E., Nathan, R. G., & Zolman, J. E. (1988). The developmental lag hypothesis in reading: Longitudinal and matched reading-level comparisons. Child Development, 59, 71–86.

Tabachnick, B.G. & Fidell, L.S. (2001). *Using Multivariate Statistics, Fourth Edition.* Needham Heights, Massachusetts: Allyn & Bacon.

Tierney, R. J., & Shanahan, T. (1991). *Research on the reading–writing relationship: Interactions, transactions, and outcomes.* Barr, Rebecca; Kamil, Michael L.; Mosenthal, Peter B.; Pearson, P. David (1991), Reprinted 1996 Handbook of reading research, Vol. 2. (pp. 246-280). Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Vellutino, F. R., Tunmer, W. E., Jaccard, J., Chen, R., & Scanlon, D. M. (2004). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. Scientific Studies of Reading, *11*(1), 3–32

Wagner, R. K., Puranik, C. S., Foorman, B., Foster, E., Wilson, L. G., Tschinkel, E., & Kantor, P. T. (2010). Modeling the development of written language. *Reading & Writing: An Interdisciplinary Journal*.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. (2010). The Test of Silent Reading Efficiency and Comprehension (TOSREC). Austin, TX: PRO-Ed.

Wilkinson, G.S. (1993) WRAT3 Wide Range Achievement Test: Administration manual (1993 ed.), Wide Range, Inc, Wilmington, DE.

Whitaker, D., Berninger, V., Johnston, J., & Swanson, L. (1994). Intraindividual differences in levels of language in intermediate grade writers: Implications for the translating process. *Learning and Individual Differences, 6,* 107–130.

Woodcock, R. (1987). *Woodcock Reading Mastery Test—Revised.* Circle Pines, MN: American Guidance Service.

Woodcock, RW & Johnson, MB *(1990) Woodcock–Johnson tests of achievement (Riverside Publishing Co, Itasca, IL), revised ed.*

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *WJIII Tests of Cognitive Abilities and Achievement.* Itasca, IL: Riverside Publishing.

BIOGRAPHICAL SKETCH

Yusra Ahmed completed her Associate of Arts degree from Florida State University, Republic of Panama, in December 2005. She pursued and completed a Bachelors degree in Psychology from Florida State University, Tallahassee, in August 2008. She worked at Florida State University as a Research Assistant with Dr. Rick Wagner and Dr. Ralph Radach. As the Bess H. Ward scholar she also completed an honors thesis in Clinical Psychology with Dr. Joyce Carbonell and Dr. Doris Gray. Yusra began her graduate studies in Developmental Psychology at Florida State University in the fall of 2008 under the advisement of Dr. Rick Wagner. Her current research interests include the co-development of reading and writing skills, and the identification and classification of learning disabilities.