Evaluating the dimensionality of first grade written composition

Young-Suk Kim[1], **Stephanie Al Otaiba**[2], **Jessica S. Folsom**[3], **Luana Greulich**[4], and **Cynthia Puranik**[5]

[1]Florida State University and Florida Center for Reading Research

[2]Southern Methodist University

[3]Florida Center for Reading Research

[4]Andrews University

[5]University of Pittsburgh

Abstract

Purpose—We examined dimensions of written composition using multiple evaluative approaches such as an adapted 6+1 trait scoring, syntactic complexity measures, and productivity measures. We further examined unique relations of oral language and literacy skills to the identified dimensions of written composition.

Method—A large sample of first grade students ($N = 527$) was assessed on their language, reading, spelling, letter writing automaticity, and writing in the spring. Data were analyzed using a latent variable approach including confirmatory factor analysis and structural equation modeling.

Results—The seven traits in the 6+1 trait system were best described as two constructs: substantive quality, and spelling and writing conventions. When the other evaluation procedures such as productivity and syntactic complexity indicators were included, four dimensions emerged: substantive quality, productivity, syntactic complexity, and spelling and writing conventions. Language and literacy predictors were differentially related to each dimension in written composition.

Conclusions—These four dimensions may be a useful guideline for evaluating developing beginning writer's compositions.

Children's writing skills, particularly at the discourse level (written composition hereafter), are typically assessed by asking children to write or compose on a given topic (e.g., the writing portion of the National Assessment of Educational Progress [NAEP], National Center for Education Statistics [NCES], 2003; Test of Written Language –4th edition, Hammill & Larsen, 2009). Various approaches to evaluation of written composition have been used, including holistic scoring, analytic scoring, quantitative scoring, and curriculum based measurement (Abbott & Berninger, 1993; Berman & Verhoevan, 2002; Mackie & Dockrell, 2004; McMaster, Du, & Pestursdottir, 2009; McMaster & Espin, 2007; Nelson, Bahr, & Van Meter, 2004; Puranik, Lombardino, & Altmann, 2008; Scott & Windsor, 2000). These various evaluation approaches differ in purposes and in the underlying

Corresponding Author: Young-Suk Kim G129 Stone Building 1114 W. Call St., Florida State University, Tallahassee, FL 32306 Phone: 850-644-0370 Fax: 850-645-8947.

assumptions about the dimensionality of written composition. It is not clear, however, how these various evaluative approaches are related, and whether they converge into a single dimension or diverge into multiple dimensions.

This question of dimensionality has not been investigated in the majority of previous studies examining written composition. Instead, a priori assumptions have been made about the dimensionality of written composition. Some research teams have presumed that written composition has a single dimension (Berninger & Abbott, 2010; NCES, 2003), or two dimensions such as quality and productivity (e.g., Abbott & Berninger, 1993; Berninger, Abbott, Abbott, Graham, & Richards, 2002; Graham, Berninger, Abbott, Abbott, & Whitaker, 1997; Olinghouse & Graham, 2009), while others theorize it has as many as seven dimensions, as represented in the 6+1 trait scoring (Northwest Regional Educational Laboratory [NREL], 2011). If indeed multiple dimensions exist for student's written composition, it is important to understand how language and literacy skills differentially predict the identified dimensions.

In the present study, we had two primary aims. The first aim was to examine how many dimensions exist when using multiple writing evaluation approaches such as adapted 6+1 traits, productivity indicators, and syntactic complexity indicators. The second aim was to examine how oral language, reading, and transcription skills (letter writing automaticity and spelling) are related to the identified dimensions of students' written composition. We addressed these aims using a large dataset of first grade students ($N = 527$) assessed at the end of the year, and employing a latent variable approach. Addressing these questions is critical to advance the field towards methods of evaluating written composition that align with theories of writing development. Furthermore, with many states now adopting the Common Core Standards that increase the prominence of standards for writing development, it is important to understand how to evaluate writing performance to inform instructional decision making.

Evaluating Written Composition

Holistic scoring, which has been used since the 1950s (Espin, Weissenburger, & Benson, 2004), implicitly assumes that written composition is unidimensional. In holistic scoring, written composition is evaluated based on the gestalt, or general impression (Wolcott & Legg, 1998) and the overall global quality of written composition is rated on a single scale (e.g., 1 to 5; Espin et al., 2004). While multiple aspects such as logical organization, development of ideas, vocabulary, spelling, and other writing conventions are considered, no specific aspect determines the overall score (see Hunter, Jones, & Randhawa, 1996) because the whole writing product is more than the sum of its parts (Myers, 1980). Holistic scoring is widely used including the NAEP writing assessment, Scholastic Aptitude Test (SAT) essay, American College Testing (ACT) essay, Graduate Record Examination (GRE) writing task, high stakes state writing tests (e.g., Florida Comprehensive Assessment Test), and in research studies (e.g., Espin, De La Paz, Scierka, & Roelofs, 2005; Olinghouse, 2008). Holistic scoring of writing evaluation varies in reliability and validity (See Espin et al., 2004 for further details), which may be partly due to variation in the weight raters place on multiple aspects considered in determining a score (e.g., development of ideas vs. spelling), particularly when scoring criteria are not clearly defined (Fitzpatrick, Ercika, Yen, & Ferrara, 1998). Furthermore, for instructional purposes, the holistic scoring approach might not be particularly useful because written compositions with the same score may have strengths and weaknesses in different aspects of the text.

Other researchers have implicitly assumed that there are two dimensions to writing, namely, quality and productivity (e.g., Abbott & Berninger, 1993; Berninger, et al., 2002; Graham et

al., 1997; Olinghouse, 2008). The quality of written composition is typically rated on content and organization of the information in written composition (e.g., Graham et al., 1997; Wechsler, 1992), whereas productivity includes the number of written words, sentences, and ideas in a composition (e.g., Abbott & Berninger, 1993; Kim et al., 2011b; Puranik et al., 2008; Wagner et al., 2011). While length, in and of itself, is certainly not the ultimate goal of writing or writing instruction, a certain level of length is necessary to elaborate and develop ideas fully with sufficient details. Studies have shown that a productivity indicator such as the total number of words in written composition is a robust measure of development (McMaster & Espin, 2007; Nelson & Van Meter, 2007; Wagner et al., 2011), and that it differentiates writing performance of school-aged children with and without language-learning disabilities (Scott & Windsor, 2000). The productivity indicators are weakly (.18 $\leq rs \leq$ .34; Wagner et al., 2011) or moderately related to writing quality (.52 $\leq rs \leq$ .60; Abbott & Berninger, 1993; Olinghouse, 2008) for students in elementary schools.

Analytic scoring allows raters to evaluate a number of predetermined aspects that are germane to good writing (Calfee & Miller, 2007; Espin et al., 2004). Diederich (1974) initially identified the following six primary aspects by analyzing "expert" reader comments on 3,557 essays: (1) ideas, which include richness, clarity, development, and relevance to topic and purpose; (2) organization, which involves the structural aspect of written composition; (3) spelling and conventions, which entails sentence structure, punctuation, and spelling; (4) wording, which includes phrasing, choice and arrangement of words; (5) style, which includes originality, individuality, and interest; and (6) handwriting. Currently, the 6+1 trait system is the most widely used analytic scale in the United States (Gansle, VanDerHeyden, Noell, Resetar, & Williams, 2006) and is used in at least 10 other countries (NREL, 2011). In this approach, written composition is rated on 5 or 6 point scales on the following predetermined seven aspects: ideas, organization, sentence fluency or flow of language, word choice, voice, writing conventions, and presentation (NREL, 2011). When using the 5 point scale, which was used in the present study, a score of 1 is descriptively defined as "experimenting", 2 "emerging", 3 "developing", 4 "capable", and 5 "experienced." While more labor intensive and costly than the holistic scoring, analytic scoring can be implemented with high reliability (Espin et al., 2004), and has shown to be moderately related to holistic scoring for children in grades 5, 8, and 11 (.40 $\leq rs \leq$ .55; Hunter et al., 1996) and fairly strongly for college students' compositions (.61 $\leq rs \leq$ .76; Freedman, 1991). Furthermore, analytic scoring may be useful for instructional purposes because it can provide a more direct link between assessment and instruction than does holistic scoring (Espin et al., 2004) because students' strengths and weaknesses in various aspects can be identified and instructionally targeted, and students' progress in response to instruction can be monitored (NREL, 2011).

Despite its wide use, empirical evidence about the factor structure of the 6+1 trait scoring is sparse. An exception is a study by Gansle et al. (2006), who used Cronbach's alpha and argued that the seven traits were unidimensional. However, Cronbach's alpha is not an ideal approach to examine dimensionality because it can be high even when measuring dissociable latent constructs (Cronbach, 1951; Green, Lissitz & Mulaik, 1977; Schmitt, 1996). An appropriate approach is the factor analytic method (Schreiber, Stage, King, Nora, & Barlow, 2006) such as confirmatory factor analysis.

Identifying the factor structure of writing is important because it would reveal the alignment of theory, assessment and evaluation, and provide a framework for developing and using specific scoring procedures for the identified writing dimensions. For instance, it has been hypothesized that written composition might be constrained by linguistic skills at multiple levels, transcription related skills, and cognitive skills (Berninger et al., 2002; Berninger & Swanson, 1994). To date, however, there is limited evidence about how these hypothesized

skills might be represented in early written composition and which language and literacy skills uniquely predict each writing dimension. Such evidence could inform writing instruction by helping teachers know which specific language and literacy skills to target in order to improve a particular dimension of written composition.

Initial efforts have been made in this line of work. Puranik et al. (2008), for instance, used an exploratory factor analysis for written compositions from fourth, fifth, and sixth grade students with language impairments, and identified three factors – productivity, syntactic complexity, and accuracy of spelling and writing conventions. Wagner and his colleagues (2011) used confirmatory factor analysis and demonstrated that first and fourth grade students' written composition was best captured by four dimensions: macro organization which included topic sentence and logical ordering of ideas; complexity which included mean length of T units and clause density; productivity which included total number of words and different words; and spelling and punctuation which included number of errors in spelling, capitalization, and punctuation.

Predictors of various dimensions of writing

The second aim of the present study was to examine relations of oral language and literacy skills to the identified writing dimensions. If written composition has multiple, dissociable dimensions (e.g., Wagner et al., 2011), then relations of oral language and literacy skills to the identified dimensions might vary. For instance, transcription skills might be more strongly related to the spelling and writing conventions dimension than to the macro organization dimension. In the present study, oral language, reading, spelling, and letter writing automaticity were included as component skills of writing based on the developmental model of writing and findings from previous studies (Berninger et al., 2002; Berninger & Swanson, 1994; Kim et al., 2011b). Oral language skills are important for writing because writing requires generation of ideas which need to be translated in language before transcribing those language representations in written texts (Berninger & Swanson, 1994). For instance, vocabulary knowledge would add interest, richness, and precision in word choice in children's written composition (Baker, Gersten, & Graham, 2003) and thus, children's oral vocabulary may be important for children's written composition. Dockrell and her colleagues (2009) showed that vocabulary was uniquely related to written composition for adolescents with specific language impairments. Children's grammatical knowledge also might contribute to their written composition because children with sophisticated grammatical knowledge are likely to express their ideas in a clear, cohesive, and accurate manner. For example, researchers have found that children with learning or language disabilities had less sophisticated vocabulary and grammatical structures than typically developing children (Fey et al., 2004; Gillam & Jonston, 1992; Scott & Windsor, 2000). Grammatical knowledge was found to be related to writing quality for typically developing third grade children (Olinghouse, 2008) and children with specific language impairments (Mackie, Dockrell, & Lindsay, in press). Oral language was also uniquely related to writing productivity for children in kindergarten (Kim et al., 2011b) and third grade (Berninger & Abbott, 2010). Thus, there is growing evidence that oral language skills may contribute to written composition.

Evidence is also accumulating for the contribution of reading skill to written composition, although the strengths of the relation vary among studies and the grade levels of participants (Berninger et al., 2002; Dockrell et al., 2009; Olinghouse & Graham, 2009; see Shanahan, 2006 for a review). Reading comprehension was related to the organization, quality, and productivity of writing (Berninger et al., 2002; Shanahan & Lomax, 1986), and to an overall writing score (Berninger & Abbott, 2010) for children in elementary and middle school grades. Similarly, children with impaired reading comprehension showed weaker story

content and organization in their writing (Cragg & Nation, 2006). Moreover, reading skill appears to be more strongly related to composition quality than to composition productivity (Abbott & Berninger, 1993; Berninger et al., 2002).

Transcription skills are necessary for writing because automatized transcription skills would free higher order cognitive resources (e.g., idea generation, revising) to be used for written composition (Abbott & Berninger, 1993; Graham, 1990; Moats, Foorman, & Taylor, 2006). Letter writing automaticity has been shown to be related to writing productivity (Berninger et al., 1997; Graham et al., 1997; Kim et al., 2011b; Wagner et al., 2011), writing quality (Graham et al., 1997; Olinghouse & Graham, 2009), macro organization (Wagner et al., 2011), and syntactic complexity (Wagner et al., 2011) for children from kindergarten to middle school. In contrast, spelling appears to be related to writing productivity (Graham et al., 1997; Graham, Harris, & Chorzempa, 2002; Kim et al., 2011b), but not to writing quality (Graham et al., 1997) for primary grade children.

To summarize, we addressed the following three research questions in the present study: (1) How many dimensions are found in children's written composition when using the adapted 6+1 trait scoring?; (2) How many dimensions are found in children's written composition when using productivity indicators and syntactic complexity indicators as well as the adapted 6+1 trait scoring?; and (3) How are oral language, reading, letter writing automaticity, and spelling skills related to the identified dimensions of written composition?

Method

Participants

Writing samples were collected from 531 first grade students (mean age = 6.61, *SD* = .39) in 34 classrooms in seven public schools in a Northern Florida community. Four writing pieces were unscorable because no words were written and thus, excluded from the statistical analysis, leaving a final sample of 527. The participants included 45% girls, 47% Black, 41% White, 2% Asian, and 10% Other, which included Hispanic, Multiracial, and unknown/ not reported. Approximately 57% of these children were eligible for free or reduced lunch. The students were participating in a larger study investigating the efficacy of core reading instruction within a response to treatment (RTI) framework (Al Otaiba et al., 2011). Writing was not the focus of the study, and there were no differences in students' writing including spelling as a function of students' condition (i.e., RTI vs. conventional approach conditions; Kim et al., 2011a).

Measures

Outcome: Written composition—Students were asked to compose a text in response to a story prompt (McMaster, Xiaquing, & Pestursdottir, 2009). This curriculum based measurement task was selected because it has been used previously with first grade students and because it is similar to state-wide curriculum-based writing assessments. The writing prompt was "One day, when I got home from school,…," which was written on the provided lined paper. Research assistants introduced the task and attempted to orient children to task expectations through the following direction. Today I'm going to ask you to write a story. Before you write, I want you to think about the story. First you will think, then you will write. You will have 30 seconds to think and 5 minutes to write. If you do not know how to spell a word, you should guess. Your story will begin with 'One day, when I got home from school.' Think of a story you are going to write that starts like that. Students were given five minutes to complete the task (McMaster et al., 2009; McMaster et al., 2011). When students copied the prompt again in their written composition, it was not included in the evaluations described below. The assessors briefly reviewed the child's written composition on the spot

after children had finished the writing task, and if the content of the written composition was not easily understandable, the assessor asked the children to read the story back to the assessor who recorded what the child meant to write.

Students' written composition was coded by the following procedures. First, the 6+1 Traits of Writing Rubric for Primary Grades (see NREL, 2011 for scoring guide and examples) was adapted. The 6+1 traits rubric includes a voice trait, but it was not included in the present study because voice was not sufficiently present for the majority of these writing samples. We also separated mechanics (capitalization and punctuation) from the spelling trait, and renamed the 'presentation' trait of the 6+1 trait to 'handwriting'. Thus, the following seven traits were evaluated in the present study: (1) Ideas for how well main ideas were developed and represented; (2) Organization for text structure; (3) Word choice for use of interesting and specific words; (4) Sentence fluency for grammatical use of sentences and flow of sentences; (5) Spelling for accuracy and for the developmental phase of spelled words; (6) Mechanics for capitalization and punctuation accuracy; and (7) Handwriting for spacing, neatness, and letter formation. These components were rated on a scale of 1 to 5. A score of zero was assigned to unscorable ones ($n = 4$). The rubric and examples used in the present study are found in the supplemental materials.

The second scoring procedure was related to writing productivity indicators such as the number of words, number of different words, and number of ideas. The number of words is a commonly used measure of compositional fluency/productivity in writing (e.g., Abbott & Berninger, 1993; Berman & Verhoevan, 2002; Mackie & Dockrell, 2004; McMaster et al., 2009; Nelson et al., 2004; Puranik et al., 2008; Scott & Windsor, 2000; Wagner et al., 2011). Words were defined as real words recognizable in the context of the child's writing despite some spelling errors (e.g., dast for best). By contrast, random strings of letters or sequences of nonsense words (e.g., EOT for 'came'; WHOWiuntris for 'in a school bus from') were not counted as words. Random strings of letters were identified by comparing a record of what the child said she had written to her written composition. These were rare and occurred less than 2% of the sample. The number of ideas was a total number of propositions, which were defined as predicate and argument. For example, "I went upstairs and took a bath" was counted as two ideas (Kim et al., 2011b). The number of different words was obtained from Systematic Analysis of Language Transcription (SALT, Miller & Chapman, 2001) after children's written composition was transcribed following the SALT guidelines.

Finally, the mean length of the T unit and clause density comprised the syntactic complexity of written composition (Wagner et al., 2010). T unit was defined as one independent clause plus any dependent clauses (Hunt, 1970; Miller & Chapman, 2001). Although the mean length of T unit in morphemes is a commonly used measure of syntactic complexity in oral language research (Paul, 2007), we used the mean length of T unit of words (total number of words divided by the total number of T units) following previous studies of written composition (Nippold, Ward-Longerman, & Fanning, 2005; Puranik et al., 2008; Scott & Windsor, 2000; Wagner et al., 2011). Clause density was calculated as a ratio of the total number of clauses divided by the total number of T units. Clause was defined as a group of words that contains a subject and a verb (Puranik et al., 2008). For instance, "I watched TV while I was eating a hot dog" was counted as two clauses.

Oral language skills—Students' oral language skills were assessed by expressive vocabulary and grammatical knowledge tasks. Vocabulary was assessed by the Picture Vocabulary subtest of Woodcock Johnson-III (WJ-III; Woodcock, McGrew, & Mather, 2001), which requires students to express the name for pictured objects. Cronbach's alpha was reported to be .70 for six-year-olds. Children's grammatical knowledge was measured by the Grammatic Completion subtest of the Test of Language Development - Intermediate,

third edition (TOLD-I: 3; Hamill & Newcomer, 1997). In this task, the student listens to a sentence read aloud with a word missing and is asked to provide grammatically correct responses for the missing part. The 28 test items include various syntactic features such as noun-verb agreement, pronoun use, plurals, and negatives. Cronbach's alpha was reported to be .91.

Reading skills—Students' reading skill was assessed by the Letter Word Identification subtest of the WJ-III (Woodcock, McGrew, & Mather, 2001), the Sight Word Efficiency of the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999), three passages of the Dynamic Indicators of Basic Early Literacy Skills Oral Reading Fluency for first grade (DIBELS – 6th edition, Good & Kaminski, 2006), and the Test of Silent Reading Efficiency and Comprehension (TOSREC; Wagner, Torgesen, Rashotte, & Pearson, 2010). TOSREC is a three min. timed test and required participants to read sentences silently and indicate whether sentences were true or false by circling 'yes' or 'no.' For example, for the statement, "An apple is blue," the correct answer is "no." There were two sample items to explain the task to the student, five practice items, and 50 test items. Total scores are calculated by counting the number of correct responses and subtracting the number of incorrect responses. The test-retest reliability was reported to be .92.

Spelling—The spelling subtest of the WJ-III (Woodcock et al., 2001) was used, which is a dictation task in which students were asked to spell words that increase in difficulty. The research assistant read each word, read the sentence with the word, and then repeated the spelling word (e.g., "dog". "I took my dog to the park." "dog"). Cronbach's alpha was reported to be .92. Traditional dichotomous (correct/incorrect) and the overall score of the spelling sensitivity score system (Apel, Masterson, & Brimo, 2011) were used as two indicators of a spelling latent variable. The overall score is coded for accuracy of entire words: Omitted words are given zero points, words that contain illegal misspellings are given one point, words that are misspelled legally are given two points, and correct spellings are given three points. The other score available in the spelling sensitivity score, the element score which assigns points for morphological and orthographic elements in a word (see Apel et al., 2011), had an extremely high correlation with the overall score ($r = .98$) and thus was not used.

Letter writing automaticity—Letter writing automaticity was assessed by asking children to write as many alphabet letters as possible in one minute with accuracy (Jones & Christensen, 1999; Kim et al., 2011b; Wagner et al., 2011). This task assessed how well children access, retrieve, and write alphabet letter forms automatically. Research assistants asked children to write all the alphabet letters in order, using lower case letters. The directions were as follows: We're going to play a game to show me how well and quickly you can write your abc's. First, you will write the lowercase of small abc's as fast and carefully as you can. Don't try to erase any of your mistakes, just cross them out and go on. When I say "ready begin", you will write the letters. Keep writing until I say stop. Ready, begin. If children finished writing 26 letters before a minute, they were asked to continue to write starting with "a" again. Children received a score for the number of correctly written letters with one point awarded for each correctly formed and sequenced letter. A 0.5 was assigned for each imprecisely formed letter (e.g., "n" must not be confusing with an "h" – it must not have a long vertical line). The following responses were scored as incorrect and earned a score of zero: (a) letters written in cursive; (b) letters written out of order; or (c) uppercase letters. The school district where the study was conducted did not teach cursive in first grade and thus, use of cursives was rare and only 3 children were found to use cursives.

Reliability

For the trait scoring of written composition, two graduate student coders were rigorously trained by the first author through a series of meetings for a total of 15 hours, and double-coded independently 15 percent of the writing sample (i.e., 78 pieces). The inter-rater agreement rates were as follows: .84 for Ideas, .92 for Organization, .84 for Word choice, .88 for Sentence fluency, .92 for Spelling, .92 for Mechanics, and .80 for Handwriting. Interrater agreement rates for the total number of words and number of ideas were .99 and .95, respectively. The number of different words was produced automatically by SALT. Transcription reliability was .95 based on 10% of written composition. The percent agreement rates in the number of T units and clauses by three graduate research assistants were .95 and .94, respectively, using 10% of written composition. For the spelling sensitivity scoring for the spelling task, percent agreement rates ranged from .92 to .99 with a master coder. For the letter writing automaticity task, percent agreement among raters was .98 for 50 children. All these reliability estimates were calculated using exact, not adjacent, agreement.

Procedures

Trained graduate students administered all the measures, which were administered in the spring (April and May). Spelling, letter writing automaticity, and writing were group administered to all consented students in their classrooms whereas reading and oral language measures were individually assessed in a quiet room at school.

Data Analysis

The primary statistical analytical method in the present study was a latent variable approach such as confirmatory factor analysis and structural equation modeling, using MPLUS 5.1 (Muthén & Muthén, 2006). Latent variable approach reduces measurement error by using common variance among multiple indicators for a latent variable instead of using a single observed variable (Bollen, 1989; Kline, 2005). Confirmatory factor analysis was used to examine dimensionality, for which alternative models were compared using a chi-square difference test between nested models (Kline, 2005; Schumacker & Lomax, 2004). Structural equation modeling was used to examine unique language and literacy predictors of the identified dimensions. Model fits were evaluated by multiple indices including chi-square, comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residuals (SRMR). Generally, RMSEA values below .085, CFI and TLI values greater than .95, and SRMR below .05 indicate an excellent model fit (Kline, 2005; Hu & Bentler, 1999). CFI and TLI values greater than .90 and SRMR less than .10 indicate a good model fit (Kline, 2005).

Results

Descriptive statistics and correlations

Descriptive statistics are presented in Table 1. The sample children's mean scores in the vocabulary and grammatical knowledge tasks were in the average range compared to the normative sample, and their literacy skills tended to be in the slightly high average range (108.01 Ms 111.73). The mean performance on the seven aspects of trait scoring ranged between 2.30 to 3.35 (emerging to developing), which is comparable to first grade students in a previous study (Gansle et al., 2006). Children wrote about 35 words and 6 ideas, on average, in their written composition. In addition, the mean length of T unit and clause density, on average, were 7.55 and 1.20, respectively for our participants; this finding is similar to the findings for first graders in Wagner et al.'s study (2011). Correlations among the 6+1 ideas, organization, word choice, sentence fluency, and spelling, and the

productivity measures were moderate, ranging from .32 to .59 (Table 2). In contrast, mechanics and handwriting were weakly related to these aforementioned variables (.09 rs .18; Table 2). Syntactic complexity measures were weakly (.12 rs .18), not related, or negatively related to many other writing measures (e.g., $rs = -.11$ & $-.13$ with organization).

Correlations among observed predictor variables such as oral language, reading, spelling, and letter writing automaticity are shown in Table 3. All the measures were moderately to strongly related to one another with an exception of the letter writing automaticity, which was weakly related to the other measures. All the correlations in Tables 2 and 3 were based on raw scores. Univariate and multivariate normality assumptions for confirmatory factor analysis were examined through skewness, kurtosis, as well as visual inspection of scatter plots and the assumptions were met.

Research Question 1: How many dimensions are found in children's written composition when using the adapted 6+1 trait scoring?—We compared two alternative CFA models: a one factor model based on the Gansle et al.'s (2006) study, and a two factor model, in which the seven aspects of the adapted trait scoring were hypothesized to capture two dimensions[1]: (1) the substantive quality dimension composed of the Ideas, Organization, Word choice, and Sentence Fluency aspects; and (2) the spelling and writing conventions dimension composed of the Spelling, Mechanics, and Handwriting aspects (Puranik et al., 2008; Wagner et al., 2011). The two factor model yielded a good model fit: χ^2 (13) = 54.05, $p < .001$; CFI = .96; TLI = .94; RMSEA = .08; and SRMR = .038. The one factor model had a statistically poorer fit to the data: χ^2 (14) = 138.47, $p < .001$; CFI = .89; TLI = .83; RMSEA = .13; and SRMR = .06. Chi-square difference statistics were significant: $\Delta \chi^2$ (1) = 84.42, $p < .001$.

Research Question 2: How many dimensions are found in children's written composition when using productivity indicators and syntactic complexity indicators as well as the adapted 6+1 trait scoring?—Confirmatory factor analyses with alternative models were conducted. That is, models with different factor structures (1, 2, 3, and 4 factors) were systematically compared and differences in model fits were examined (see Table 4). A model in which all the indicators were assumed to capture a single dimension yielded a poor model fit (see Model 6). Two alternative models of two latent variables were fitted (Models 4 and 5). In Model 4, indicators of substantive quality and productivity were forced to form one factor and the indicators of writing conventions and syntactic complexity to form the other factor. In Model 5, indicators of substantive quality and syntactic complexity were forced to form one factor and the indicators of writing conventions and productivity to form the other factor. Neither model yielded a good fit to the data. Competing 3 factor models were also fitted, but the model fits were not good (Table 4). The four factor model had the best fit to the data. The chi-square differences between the four factor model (Model 1) and the other alternative models were all statistically significant: $\Delta \chi^2$ 353.79; 2 Δdf 5; $ps < .001$. These results suggest that the four factors, namely the substantive quality, productivity, syntactic complexity, and spelling and writing conventions, capture the data well, and therefore, these four dimensions were used as the outcomes in subsequent structural equation models.

Research Question 3: How are letter writing automaticity, spelling, reading, and oral language skills related to the identified dimensions of written composition?—Before fitting structural equation models, measurement models were

[1]As preliminary analysis, exploratory factor analysis was also conducted which suggested the two factor structure used in the confirmatory factor analysis.

constructed for the language and literacy variables, and were all appropriate (see factor loadings in Table 1). For the reading latent variable, residuals were allowed to co-vary between WJ-III Letter-Word Identification and TOWRE Sight Word Efficiency, given that both assess lexical or word level reading skills. Correlations among language, literacy, writing dimensions latent variables, and the observed letter writing automaticity variable are presented in Table 5. The majority of the variables were moderately to strongly related, but the syntactic complexity dimension was weakly related to all the other variables (.02 ≤ rs ≤ . 14).

Next, structural equation models were fitted to examine unique relations of oral language, reading, spelling, and letter writing automaticity to each of the four outcomes of written composition. Given multiple outcomes, Benjamini-Horchberg corrections (What Works Clearing House Procedures and Standards Handbook, 2012) were applied to determine statistical significance. The model fit for the substantive quality outcome was good: χ^2 (80) = 349.45, $p < .001$; CFI = .97; TLI = .96; RMSEA = .08; and SRMR = .04. Thirty five percent of the total variance in the substantive quality outcome was explained by the language and literacy predictors. Oral language, reading, and letter writing automaticity (.14 ≤ βs ≤ .30) were uniquely and positively related to the substantive quality of written composition whereas spelling was not ($\beta = -.001$; see Figure 1a). The model fit for the productivity outcome was good: χ^2 (67) = 360.97, $p < .001$; CFI = .97; TLI = .95; RMSEA = .09; and SRMR = .04. Twenty nine percent of the total variance in the productivity outcome was explained by the language and literacy predictors. As shown in Figure 1b, reading ($\beta = .40$) and letter writing automaticity ($\beta = .24$) were uniquely and positively related to the productivity of written composition whereas oral language ($\beta = -.02$) and spelling ($\beta = .02$) were not. The model fit for the syntactic complexity outcome was good: χ^2 (56) = 300.98, $p < .001$; CFI = .97; TLI = .95; RMSEA = .09; and SRMR = .04. However, only 3% of the total variance in the syntactic complexity outcome was explained by the included predictors. As shown in Figure 1c, spelling ($\beta = .16$) was uniquely but weakly related to syntactic complexity whereas oral language, reading, and letter writing automaticity were not. Finally, the model fit for the spelling and writing conventions outcome was good: χ^2 (67) = 409.19, $p < .001$; CFI = .96; TLI = .95; RMSEA = .10; and SRMR = .046. Fifty eight percent of total variance in the spelling and writing conventions outcome was explained by the language and literacy predictors. As shown in Figure 1d, the spelling predictor was strongly related to the spelling and writing conventions outcome of written composition ($\beta = .62$). Reading and letter writing automaticity were positively, but weakly related to the spelling and writing conventions outcome ($\beta s = .16$ & .10, respectively) whereas oral language ($\beta = -.06$) was not.

Discussion

The present study showed that when using the adapted seven aspects of the 6+1 trait scoring, written composition was comprised of two dimensions – substantive quality, and spelling and writing conventions. These findings suggest that the seven aspects of trait scoring are not unidimensional (Gansle et al., 2006); nor are they seven unique dimensions, at least for the adapted version used in the present study and for first grade children. Instead, the extent to which ideas are developed, organized, and expressed using interesting and descriptive words and appropriate grammatical structures appears to capture a single dimension of substantive quality while aspects related to spelling and handwriting capture another dissociable dimension. These findings, however, should not be taken to negate utility of the seven aspects of 6+1 trait scoring for evaluation and instructional purposes. While ideas, organization, sentence fluency, and word choice share enough common variance to be described as a single dimension, it might be useful to focus on each of these individual aspects instructionally. It should be noted that children might have avoided words that are

challenging to spell, which may have limited sophisticated vocabulary use in written composition. Further exploration is needed regarding the extent of the avoidance, and its impact on the factor structure of written composition.

When other frequently used quantitative indices were considered in addition to the 6+1 trait scoring, the following four emerged as separable dimensions: substantive quality, syntactic complexity, productivity, and spelling and writing conventions. These findings expand previous studies (Puranik et al., 2008; Wagner et al., 2011) by using a widely used 6+1 trait scoring and a large sample of first grade students. These four dimensions partially support a developmental model of writing (Berninger & Swanson, 1994; Berninger et al., 2002), which proposed that language skills at the discourse, sentence, and word levels as well as transcription skills can constrain or support writing development. The syntactic complexity dimension in the present study may correspond to the sentence level language skill, and the productivity dimension may correspond to the word level language skill. In addition, spelling and writing conventions align closely with transcription skills (Berninger & Swanson, 1994). The substantive quality dimension appears to capture not only ideas and organization, but also children's word choice and sentence uses. This result might suggest that some aspects of lexical and sentence level skills are important foundations for discourse level skills such that they add to the overall quality of written composition. An alternative explanation is a method effect. The present findings might be attributed to the fact that children's vocabulary use and sentence level skills were coded on a rating scale using the 6+1 trait scoring as were the other aspects such as idea development and organization. The question remains whether results would be different if vocabulary use and sentence level skills were evaluated in a different manner than how ideas and organization were evaluated. For instance, vocabulary use has been examined by counting the number of multisyllabic words (7 or longer words; Olinghouse & Leaird, 2009); counting literate vocabulary and connectives (Nippold et al., 2005); and counting different types of subordinate clauses such as relative, adverbial and nominal (Nippold et al., 2005). Recent data with Korean-speaking children suggest that the number of academic vocabulary and connectives in written composition is best considered as part of substantive quality rather than a dissociable dimension (Kim, Park, & Park, in press).

The present study also expanded our understanding of writing skills by examining how oral language and literacy skills are related to the identified multiple dimensions of written composition. Previous studies tended to examine either a single dimension (Berninger & Abbott, 2010; Kim et al., 2011b) or two dimensions of writing productivity and quality (e.g., Abbott & Berninger, 1993; Berninger & Swanson, 1994; Graham et al., 1997; Olinghouse, 2008). To our knowledge, this is the first study to investigate the unique relations of oral language and literacy skills to the systematically identified multiple dimensions of written composition. Our findings showed that the majority of the language and literacy predictors were moderately related to the four dimensions of written composition in bivariate examinations. When accounting for other skills, children's oral language was uniquely related to the substantive quality dimension. This provides further support for the role of oral language skills in children's writing development (Berninger & Abbott, 2010; Kim et al., 2011b; Shanahan, 2006), but expands previous studies by showing a specific dimension of written composition which is uniquely influenced by children's oral language skills.

Children's reading ability was uniquely related to the substantive quality, productivity, and spelling and writing conventions dimensions after accounting for oral language, spelling, and letter writing automaticity. The unique contribution of children's reading skill to the substantive quality dimension might be due to fact that the reading latent variable in the present study included not only word reading skills, but also sentence and connected text level reading skills such as reading fluency. Children's reading skills in both sentences and

connected text might help their written expression in terms of logical representation of ideas, and sentence level constructions. Additionally, children with higher reading skills tended to write longer written compositions even after accounting for oral language and transcription skills. Children with higher reading skills may read more, which may influence interest and ability to generate more text in written composition (Berninger et al., 2002). Overall, these results are convergent with previous studies on the relation between reading and writing (Abbott & Berninger, 1993; Berninger et al., 2002; Shanahan & Lomax, 1986).

Transcription skills such as spelling and letter writing automaticity were associated with multiple dimensions of written composition for beginning writers. Letter writing automaticity was uniquely related to three dimensions, but not to syntactic complexity. This finding extends previous studies (e.g., Berninger et al., 1996; Graham et al., 1997; Kim et al., 2011b; Wagner et al., 2011) by examining the independent contribution of letter writing automaticity to multiple dimensions of written composition after accounting for the other oral language and literacy skills. Taken together with findings from previous studies, these results suggest that achieving automaticity in writing letters is an important prerequisite skill that supports various dimensions of written composition for beginning writers.

Another transcription skill, spelling, was uniquely and strongly related to the spelling and writing conventions dimension of written composition after accounting for the other language and literacy predictors. It was somewhat surprising that spelling was not uniquely related to the productivity dimension because spelling would be expected to constrain the amount of text one writes particularly in the beginning phase of writing development (Graham et al., 1997; Kim et al., 2011b). The lack of a unique relation in the present study might also be due to a high correlation of spelling with reading at this point of development ($r = .79$). Interestingly, spelling skill was the only unique predictor that was related to the syntactic complexity dimension of written composition, suggesting that the extent of producing syntactically complex sentence is constrained by spelling skill at this beginning stage of writing development. The lack of a relation of oral language to the syntactic complexity dimension was somewhat surprising. Syntactic development is generally marked by an increase in sentence length and the use of morphological markers in oral language (Nippold et al., 2005; Nippold, 2010; Nelson & van Meter, 2007) and has been examined in children's written compositions (e.g., Dockrell et al., 2009; Nippold et al., 2005; Scott & Windsor, 2000; Wagner et al., 2011). Therefore, one might expect vocabulary and grammar knowledge of the oral language latent variable to be related to the syntactic complexity dimension. One explanation for the lack of such a relation might be that the oral language tasks in the present study measured a somewhat different aspect of sentence level skills. That is, providing a grammatically correct form of a given word as measured by TOLD Gramatic Completion may not be predictive of syntactic complexity as captured by clause density and mean length of T units in written composition.

Limitations and Directions for Future Research

Several limitations of the present study and directions future studies are important to note here. First, the findings of the present study are from the end of first grade when the majority of children are in the beginning phase of writing development. Thus, these relations cannot be generalized to later phases of development. We are currently conducting longitudinal studies to investigate dimensions and factor structures of written composition for children in different developmental phases, and examine whether relative weights of various language and cognitive skills vary along a developmental continuum. Second, only one writing prompt was used to assess children's writing skill. Studies have shown that writing quality in different topics and genres varied for students in grades two and twelve (see Graham, Harris, & Hebert, 2011 for a review). A future study needs to replicate the present study

using multiple writing prompts of different genres and different types of writing such as a journal. Third, the oral language latent variable did not include discourse level skills. Future studies should include more comprehensive oral language skills including both comprehension and expression (e.g., Berninger & Abbott, 2010). Furthermore, other indicators of syntax in written composition should be explored. Although clause density in written composition has been shown to increase until eighth grade for typically developing children (Nippold et al., 2005) and mean length of T units is a used measure of syntactic complexity (Nippold et al., 2005; Puranik et al., 2005; Wagner et al., 2011), these capture the complexity aspect of syntax, but not knowledge of grammaticality. Thus, other indicators such as use of cohesive devices and syntactic structures and diversity of syntactic structures (e.g., Nippold et al., 2005) might capture the syntactic dimension in children's written composition. A potential source to capture such indicators might be several variables in Coh-Metrix (Graesser, McNamara, & Kulikowich, 2011) as explored by Puranik, Wagner, Kim, and Lopez (2012). Finally, the overall amount of total variance explained by the included oral language and literacy predictors was fairly limited, similar to a previous study with kindergartners (Kim et al., 2011). Thus, other potential factors to be explored include cognitive, metacognitive, affective factors, self-regulation, attentional difficulties, and instruction (Berninger & Winn, 2006; Kim, Al Otaiba, Sidler, & Greulich, 2013).

In conclusion, the present study revealed four dimensions of students' written composition for first grade students: substantive quality, syntactic complexity, productivity, and spelling and writing conventions. Evaluation of students' written expression may focus on these four dimensions. Furthermore, children's ability in oral language, reading, spelling, and letter writing automaticity are differentially related to these various dimensions. Thus, instructional attention to these language and literacy skills will be important to facilitate development of writing skills for children.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Abbott RD, Berninger VW. Structural equation modeling of relationships among development skills and writing skills in primary- and intermediate-grade writers. Journal of Educational Psychology. 1993; 85:478–508.

Al Otaiba S, Connor CM, Folsom JS, Greulich L, Meadows J, Li Zhi. Assessment data-informed guidance to individualize kindergarten reading instruction: Findings from a cluster-randomized control field trial. Elementary School Journal. 2011; 111:535–560. [PubMed: 21818158]

Apel, K.; Masterson, JJ.; Brimo, D. Spelling assessment and intervention: A multilinguistic approach to improving literacy outcomes. In: Kamhi, AG.; Catts, HW., editors. Language and reading disabilities. 3rd ed. Pearson; Boston, MA: 2011. p. 226-243.

Baker S, Gersten R, Graham S. Teaching expressive writing to students with learning disabilities: Research-based applications and examples. Journal of Learning Disabilities. 2003; 36:109–123. [PubMed: 15493427]

Berman R, Verhoevan L. Cross-linguistic perspectives on the development of text-production abilities. Written Language and Literacy. 2002; 5:1–43.

Berninger VW, Abbott RD. Listening comprehension, oral expression, reading comprehension, and written expression: Related yet unique language systems in grades 1, 3, 5, and 7. Journal of Educational Psychology. 2010; 102:635–651. [PubMed: 21461140]

Berninger VW, Abbott RD, Abbott SP, Graham S, Richards T. Writing and reading: Connections between language by hand and language by eye. Journal of Learning Disabilities. 2002; 35:39–56. [PubMed: 15490899]

Berninger, VW.; Swanson, HL. Children's writing; toward a process theory of the development of skilled writing. In: Butterfield, E., editor. Children's writing: Toward a process theory of development of skilled writing. JAI Press; Greenwich, CT: 1994. p. 57-81.

Berninger VW, Vaughan KB, Graham S, Abbott RD, Abbott SP, Rogan LW, Brooks A, Reed E. Treatment of handwriting problems in beginning writers: Transfer from handwriting to composition. Journal of Educational Psychology. 1997; 89:652–666.

Berninger, VW.; Winn, W. Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In: MacArthur, CA.; Graham, S.; Fitzgerald, J., editors. Handbook of writing research. Guilford Press; New York, NY: 2006. p. 96-114.

Bollen, KA. Structural equations with latent variables. John Wiley & Sons; New York, NY: 1989.

Calfee, RC.; Miller, RG. Best practices in writing assessment. In: Graham, S.; MacArthur, C.; Fitzgerald, J., editors. Best Practices in Writing Instruction. Guilford Press; New York, NY: 2007. p. 265-286.

Cragg L, Nation K. Exploring written narrative in children with poor reading comprehension. Educational Psychology. 2006; 26:55–72.

Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951; 16:297–334.

Diederich, PB. Measuring growth in English. National Council of Teachers of English; Urbana, IL: 1974.

Dockrell J, Lindsay G, Connelly V. The impact of specific language impairment on adolescents' written text. Exceptional Children. 2009; 75:427–446.

Espin CA, De La Paz S, Scierka BJ, Roelofs L. Relation between curriculum-based measures in written expression and quality and completeness of expository writing for middle-school students. Journal of Special Education. 2005; 38:208–217.

Espin CA, Weissenburger JW, Benson BJ. Assessing the writing performance of students in special education. Exceptionality. 2004; 12:55–66.

Fey M, Catts H, Proctor-Williams K, Tomblin B, Zhang X. Oral and written story composition skills of children with language impairment. Journal of Speech, Language, and Hearing Research. 2004; 47:1301–1318.

Fitzpatrick AR, Ercikan K, Yen WM, Ferrara S. The consistency between rater scoring in different test years. Applied Measurement in Education. 1998; 11:195–208.

Freedman, SW. Evaluating writing: Linking large-scale testing and classroom assessment (Occasional Paper No. 27). University of California, Center for the Study of Writing; Berkeley: 1991.

Gansle KA, VanDerHeyden AM, Noell GH, Resetar JL, Williams KL. The technical adequacy of curriculum-based and rating-based measures of written expression for elementary school students. School Psychology Review. 2006; 35:435–450.

Gillam R, Jonston J. Spoken and written language relationships in language/learning-impaired and normally achieving school-age children. Journal of Speech and Hearing Research. 1992; 35:1303–1315. [PubMed: 1283612]

Good, RH.; Kaminski, RA., editors. Dynamic Indicators of Basic Early Literacy Skills. 6th ed. Institute for the Development of Educational Achievement; Eugene, OR: 2007.

Graham S. The role of production factors in learning disabled students' compositions. Journal of Educational Psychology. 1990; 82:781–791.

Graham S, Berninger VW, Abbott RD, Abbott SP, Whitaker D. Role of mechanics in composing of elementary school students: A new methodological approach. Journal of Educational Psychology. 1997; 89:170–182.

Graham S, Harris KR, Chorzempa BF. Contribution of spelling instruction to the spelling, writing, and reading of poor spellers. Journal of Educational Psychology. 2002; 94:669–686.

Graham, S.; Hebert, M. Writing to read: The evidence-base for how writing can improve reading. Alliance for Excellent Education; Washington, DC: 2010.

Graham, S.; Harris, K.; Hebert, MA. A Carnegie Corporation Time to Act report. Alliance for Excellent Education; Washington, DC: 2011. Informing writing: The benefits of formative assessment.

Graesser AC, McNamara DS, Kulikowich JM. Coh-Metrix: Providing multilevel analyses of text characteristics. Educational Researcher. 2011; 40:223–234.

Green SB, Lissitz RW, Mulaik SA. Limitations of coefficient alpha as an index of test unidimensionality. Educational and Psychological Measurement. 1977; 37:827–838.

Hammill, DD.; Larsen, SC. Test of Written Language. 4th edition. Pro-ed; Austin, TX: 2009.

Hamill, DD.; Newcomer, PL. Test of Language Development-Intermediate. 3rd edition. Pro-Ed; Austin, TX: 1997. (TOLD-I: 3)

Hu L-T, Bentler PM. Cutoff criteria for fit indices in covariance structure analysis. Conventional criteria versus new alternatives. Structural Equation Modeling. 1999; 6:1–55.

Hunt, K. Grammatical structures written at three grade levels (Research Report No. 3.). National Council of Teachers of English; Champaign, IL: 1965.

Hunter DM, Jones R, Randhawa BS. The use of holistic versus analytic scoring for large-scale assessment of writing. The Canadian Journal of Program Evaluation. 1996; 11:61–85.

Jones D, Christensen CA. Relationship between automaticity in handwriting and students' ability to generate written text. Journal of Educational Psychology. 1999; 91:44–49.

Kim, Y-S.; Al Otaiba, S.; Folsom, JS.; Greulich, L.; Puranik, C. Measuring quality of writing for beginning writers. Paper presented at the Society for Scientific Studies of Reading; FL, US. 2011a.

Kim Y-S, Al Otaiba S, Puranik C, Sidler JF, Gruelich L, Wagner RK. Componential skills of beginning writing: An exploratory study at the end of kindergarten. Learning and Individual Differences. 2011b; 21:517–525. [PubMed: 22267897]

Kim Y-S, Al Otaiba S, Silder JF, Greulich L. Language, literacy, attentional behaviors, and instructional quality predictors of written composition for first graders. Early Childhood Research Quarterly. 2013; 28:461–469. [PubMed: 24062600]

Kim Y-S, Park C, Park Y. Is academic language use a separate dimension in beginning writing? Evidence from Korean children. Learning and Individual Differences. (in press).

Kim Y-S, Wagner RK, Foster L. Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of first-grade readers. Scientific Studies of Reading. 2011; 15:338–362. [PubMed: 21747658]

Kline, B. Principles and practice of structural equation modeling. Second edition. Guilford Press; New York, NY: 2005.

Kuhn MR, Schwanenflugel PJ, Meisinger EB. Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. Reading Research Quarterly. 2010; 45:232–253.

Mackie C, Dockrell J. The nature of written language deficits in children with SLI. Journal of Speech, Language, and Hearing Research. 2004; 47:1469–1483.

Mackie C, Dockrell G, Lindsay G. An evaluation of the written texts of children with SLI: The contributions of oral language, reading and phonological short-term memory. Reading and Writing: An Interdisciplinary Journal. (in press).

McMaster KL, Du X, Pestursdottir A. Technical features of curriculum-based measures for beginning writers. Journal of Learning Disabilities. 2009; 42:41–60. [PubMed: 19103799]

McMaster K, Espin C. Technical features of curriculum-based measures in writing: A literature review. Journal of Special Education. 2007; 41:68–85.

McMaster KL, Xiaoqing D, Pestursdottir AL. Technical features of curriculum-based measures for beginning writers. Journal of Learning Disabilities. 2009; 42:41–60. [PubMed: 19103799]

McMaster KL, Du X, Yeo S, Deno SL, Parker D, Ellis T. Curriculum-based measures of beginning writing: Technical features of the slope. Exceptional Children. 2011; 77:185–206.

Miller, J.; Chapman, R. Systematic analysis of language transcripts. Version 7.0. Waisman Center, University of Wisconsin-Madison; Madison, WI: 2001. [computer software]

Moats LC, Foorman BR, Taylor WP. How quality of writing instruction impacts high-risk fourth graders' writing. Reading and Writing: An Interdisciplinary Journal. 2006; 19:363–391.

Muthén, LK.; Muthén, BO. Mplus. Muthén and Muthén; Los Angeles: 2006.

National Center for Education Statistics. Persky, HR.; Dane, MC.; Jin, Y. [Retrieved May 23, 2010] The Nation's Report Card: Writing 2002, NC ES 2003-529. 2003. http://nces.ed.gov/

Nelson, NW.; Bahr, CM.; Van Meter, AM. The writing lab approach to language instruction and intervention. Paul H. Brookes; Baltimore, MD: 2004.

Nelson NW, Van Meter AM. Measuring written language ability in narrative samples. Reading and Writing Quarterly. 2007; 23:287–309.

Nippold, MA. Language sampling with adolescents. Plural Publishing; San Diego, CA: 2010.

Nippold MA, Ward-Longerman JM, Fanning JL. Persuasive writing in children, adolescents, and adults: A study of syntactic, semantic, and pragmatic development. Language, Speech, and Hearing Services in Schools. 2005; 36:125–138.

Northwest Regional Educational Laboratory. 6+1 Trait® Writing. 2011. Retrieved from http://educationnorthwest.org/traits

Olinghouse NG. Student- and instruction-level predictors of narrative writing in third-grade students. Reading and Writing: An Interdisciplinary Journal. 2008; 21:3–26.

Olinghouse NG, Graham S. The relationship between discourse knowledge and the writing performance of elementary-grade students. Journal of Educational Psychology. 2009; 101:37–50.

Olinghouse NG, Leaird JT. The relationship between measures of vocabulary and narrative writing quality in second and fourth-grade students. Reading and Writing: An Interdisciplinary Journal. 2009; 22:54–565.

Paul, R. Language disorders from infancy through adolescence. 3rd ed. Mosby, Inc; St. Louis: 2007.

Puranik C, Lombardino L, Altmann L. Assessing the microstructure of written language using a retelling paradigm. American Journal of Speech Language Pathololgy. 2008; 17:107–120.

Puranik, CS.; Wagner, RK.; Kim, Y-S.; Lopez, D. Assessing elementary students' transcription and text generation during written translation: A multivariate approach. In: Fayol, M.; Alamargot; Berninger, V., editors. Translation of Thought to Written Text While Composing: Advancing Theory, Knowledge, Methods, and Applications. Psychology Press/Taylor; New York, NY: 2012.

Scott C, Windsor J. General language performance measures in spoken and written discourse produced by school-age children with and without language learning disabilities. Journal of Speech, Language, and Hearing Research. 2000; 43:324–339.

Schmitt N. Uses and abuses of coefficient alpha. Psychological Assessment. 1996; 8:350–353.

Schreiber JB, Stage FK, King J, Nora A, Barlow EA. Reporting structural equation modeling and confirmatory factor analysis: A review. Journal of Educational Research. 2006; 99:323–337.

Schumacker, RE.; Lomax, RG. A beginner's guide to structural equation modeling. 2nd dd. Lawrence Erlbaum Associates; Mahwah, NJ: 2004.

Shanahan, T. Relations among oral language, reading, and writing development. In: MacArthur, CA.; Graham, S., et al., editors. Handbook of writing. Guilford Press; New York, NY: 2006. p. 171-183.

Shanahan T, Lomax RG. An analysis and comparison of theoretical models of the reading-writing relationship. Journal of Educational Psychology. 2006; 78:116–123.

Torgesen, JK.; Wagner, RK.; Rashotte, CA. Test of Word Reading Efficiency. PRO-ED; Austin, TX: 1999.

Wagner RK, Puranik CS, Foorman B, Foster E, Tschinkel E, Kantor PT. Modeling the development of written language. Reading and Writing: An Interdisciplinary Journal. 2011; 24:203–220.

Wagner, RK.; Torgesen, J.; Rashotte, CA.; Pearson, N. Test of Sentence Reading Efficiency and Comprehension. Pro-Ed; Austin, TX: 2010.

Wechsler, D. Wechsler Individual Achievement Test. Psychological Corporation; Orlando, FL: 1992.

What Works Clearing House Procedures and Standards Handbook. Version 2.12012. Downloaded from http://ies.ed.gov/ncee/wWc/pdf/reference_resources/wwc_procedures_v2_1_standards_handbook.pdf

Wolcott, W.; Legg, SM. An overview of writing assessment: Theory, research, and practice. National Council of Teachers of English; Urbana, IL: 1998.

Woodcock, RW.; McGrew, KS.; Mather, N. Woodcock-Johnson III Tests of Achievement. Riverside Publishing; Itasca, IL: 2001.

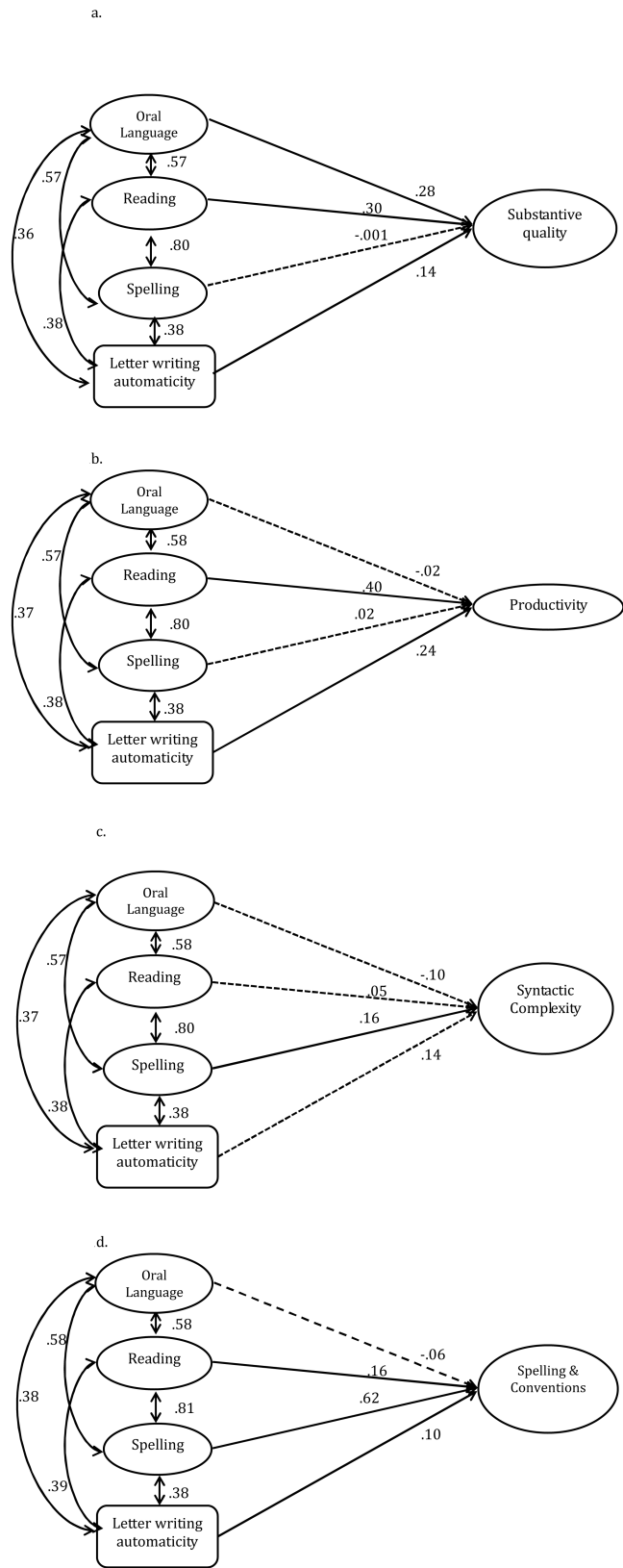**Figure 1a, 1b, 1c, and 1d.**
Standardized structural regression weights among oral language, reading, spelling, letter writing automaticity for four dimensions of written composition: substantive quality (1a); productivity (1b); syntactic complexity (1c); and spelling and writing conventions (1d). Solid lines represent statistically significant paths and dashed lines statistically nonsignficant paths.

Table 1

Descriptive statistics and standardized loadings and residuals with their respective standard error (SE)

	M (SD)	Min – Max	Loading (SE)	Residuals(SE)
Writing substantive quality				
Trait: Ideas	3.05 (.71)	0 – 4	.71 (.03)	.50 (.04)
Trait: Organization	2.83 (.65)	0 – 4	.72 (.03)	.49 (.04)
Trait: Word Choice	2.55 (.72)	0 – 4	.83 (.02)	.32 (.03)
Trait: Sentence Fluency	2.86 (.62)	0 – 4	.57 (.03)	.67 (.04)
Spelling and writing conventions				
Trait: Spelling	3.35 (.83)	0 – 5	.82 (.03)	.33 (.04)
Trait: Mechanics	2.30 (1.02)	0 – 5	.55 (.04)	.71 (.04)
Trait: Handwriting	2.71 (.87)	0 – 5	.53 (.04)	.72 (.04)
Writing Productivity				
Number of words	34.61 (15.00)	1 - 87	.99 (.004)	.02 (.008)
Number of different words	23.52 (8.76)	1 - 53	.93 (.007)	.13 (.01)
Number of ideas	6.28 (2.73)	0 – 16	.90 (.009)	.19 (.02)
Writing Syntactic complexity				
Mean length T unit	7.55 (2.58)	1 - 27	.68 (.02)	.00 ([*])
Clause density	1.20 (.30)	0 – 4	1.00 (.00)	.54 (.03)
Oral language				
TOLD Grammatic Completion	17.26 (6.28)	0 – 28	.80 (.03)	.35 (.05)
TOLD Grammatic Completion –SS	9.37 (3.09)			
WJ Vocabulary	20.17 (3.30)	11 - 29	.73 (.03)	.47 (.05)
WJ Vocabulary - SS	100.33 (11.10)			
Reading				
WJ Letter Word ID	38.95 (7.92)	15 – 62	.89 (.01)	.21 (.02)
WJ Letter Word ID - SS	111.73 (13.21)		---	---
TOWRE SWE	46.29 (16.10)	3 – 81	.93 (.007)	.14 (.01)
TOWRE SWE - SS	108.76 (14.65)		---	---
ORF passage 1	78.26 (44.72)	3 – 242	.98 (.002)	.04 (.004)
ORF passage 2	75.80 (44.68)	0 – 223	.98 (.002)	.03 (.004)
ORF passage 3	70.84 (43.35)	0 – 227	.96 (.004)	.07 (.007)
TOSREC	24.35 (9.16)	2 – 50	.83 (.01)	.31 (.02)
TOSREC - SS	108.01 (12.98)		---	---
Spelling				
WJ Spelling	25.20 (5.27)	7 – 45	.99 (.004)	.03 (.008)
WJ Spelling - SS	108.67 (15.57)		---	---
Spelling SSS score (overall)	42.51 (17.87)	0 - 107	.98 (.004)	.04 (.008)
Letter writing automaticity	17.96 (7.40)	0 – 44	---	---

Note: SS = Standard Score; TOLD = Test of Language Development; WJ Letter Word ID = Woodcock Johnson Letter Word Identification; TOWRE SWE = Test of Word Reading Efficiency: Sight Word Efficiency; ORF = Oral reading fluency; TOSREC = Test of Sentence Reading Efficiency and Comprehension; SSS = Spelling Sensitivity Score. All loadings are statistically significant at .05 level.

*Not estimated because it was set at zero based on preliminiary analysis.

Table 2

Correlations among observed variables for students' written composition

	1	2	3	4	5	6	7	8	9	10	11
1. Ideas	---										
2. Organization	.56***	---									
3. Word choice	.48***	.39***	---								
4. Sentence fluency	.55***	.59***	.45***	---							
5. Spelling	.38***	.43***	.32***	.48***	---						
6. Mechanics	.22***	.31***	.28***	.39***	.39***	---					
7. Handwriting	.19***	.20***	.18***	.30***	.48***	.36***	---				
8. Number of words	.44***	.40***	.34***	.55***	.34***	.10*	.11*	---			
9. Number of different words	.47***	.46***	.42***	.58***	.39***	.18***	.15***	.92***	---		
10. Number of ideas	.44***	.40***	.33***	.56***	.29***	.09*	.10*	.89***	.82***	---	
11. Mean length of T units	−.03	−.11**	.08	−.03	.12**	−.03	.06	.14**	.18***	−.03	---
12. Clause density	−.03	−.13**	.002	−.03	.04	−.02	.06	.05	.08	−.03	.68***

*
$p < .05$,

**
$p < .01$,

***
$p < .001$

Table 3

Correlations among observed variables for language, reading, spelling, and letter writing automaticity

	1	2	3	4	5	6	7	8	9	10
1. Grammatic Completion	---									
2. WJ Picture Vocabulary	.59	---								
3. WJ LWID	.51	.46	---							
4. Sight Word Efficiency	.44	.37	.88	---						
5. ORF passage 1	.46	.40	.86	.91	---					
6. ORF passage 2	.46	.40	.87	.91	.96	---				
7. ORF passage 3	.45	.40	.85	.88	.95	.95	---			
8. TOSREC	.43	.42	.72	.78	.82	.81	.80	---		
9. WJ Spelling	.46	.39	.86	.78	.76	.78	.76	.66	---	
10. SSS overall score	.49	.41	.86	.78	.76	.77	.76	.65	.97	---
11. Letter writing automaticity	.31	.26	.37	.36	.37	.38	.36	.38	.37	.37

Note: All correlation coefficients are statistically significant at .001 level.

LWID = Letter Word Identification

ORF = Oral reading fluency

TOSREC = Test of Sentence Reading Efficiency and Comprehension

SSS = Spelling Sensitivity Score

Table 4

Model fits for various alternative models for the writing composition outcome variables

Model		$\chi^2(df)$	CFI (TLI)	RMSEA	SRMR
1	4 latent variables (substantive quality; productivity; spell; complexity)	293.38 (49)	.93 (.91)	.097	.049
2	3 latent variables (Substantive quality & productivity; Spell; Complexity)	776.20 (51)	.80 (.75)	.16	.11
3	3 latent variables (Substantive quality & Complexity; Spell; productivity)	647.17 (51)	.84 (.79)	.15	.089
4	2 latent variables (Substantive quality & productivity; Spelling & complexity)	1101.45 (53)	.72 (.65)	.19	.13
5	2 latent variables (Substantive quality & complexity; Spelling & productivity)	984.30 (53)	.75 (.69)	.18	.135
6	1 latent variables	1306.60 (54)	.66 (.59)	.21	.21

Note: The chi-square differences between Model 1 and the other alternative models were all statistically significant: $\Delta\chi^2$ 353.79; 2 Δdf 5; *ps* < .001

Table 5

Correlations among latent variables and with letter writing automaticity observed variable

	1	2	3	4	5	6	7
1. W Substantive quality	---						
2. W Productivity	.65	---					
3. W Syntactic complexity	−.05	.13	---				
4. W Spelling and writing conventions	.65	.37	.12	---			
5. Language	.50	.21	.02	.40	---		
6. Reading	.51	.44	.13	.69	.58	---	
7. Spelling	.44	.34	.15	.78	.56	.79	---
8. Letter writing automaticity	.36	.34	.06	.39	.37	.38	.38

Note: Coefficients equal to or greater than .12 are statistically significant at .05 level.

W = Written composition