An Analysis of First-Grade Writing Profiles and their Relationship to Compositional Quality

David L. Coker, Jr.[a]

Kristen D. Ritchey[a]

Ximena Uribe-Zarain[b]

Austin S. Jennings[a]

[a] University of Delaware

[b]Missouri State University

Author Note

Published in *Journal of Learning Disabilities*, 2018

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A110484 to the University of Delaware. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Abstract

In order to help all students meet the writing expectations of the Common Core State Standards (National Governors Association Center for Best Practices, 2010), researchers need a deeper understanding of the characteristics of struggling writers. The purpose of this study was to explore the profiles of students who may have or be at risk for writing disabilities. First-grade students (N = 391) were assessed using three writing assessments (spelling, sentence writing fluency, writing achievement) at the end of the school year. Using latent profile analysis, students were identified as fitting into one of five profiles (At Risk, Low Fluency, Low Writing, Average, and Above Average). Students also wrote narrative and descriptive texts that were scored multiple ways. Using confirmatory factor analysis, four common factors were identified: Quality/Length, Spelling, Mechanics and Syntax. Students in the At Risk profile wrote narratives and descriptions that scored lower on all aspects of writing when compared to students in the Average and Above Average profiles. These findings provide further evidence of the distinct difference among writers as early as first grade, and they offer insight into the characteristics of at-risk writers. The implications of these findings for instruction and assessment and directions for future research are described.

Researchers interested in learning disabilities have focused the majority of their attention on reading disabilities, but writing disabilities represent an important and underexplored area for research. As many as 40 percent of students who receive special education services under the Individuals with Disabilities Education Improvement Act (OSEP, 2014) are identified with specific learning disabilities. This statistic is not disaggregated by type of learning disability (reading, written language, mathematics, or multiple academic areas), which makes it difficult to have a clear picture of how many students experience writing disabilities. However, Katusic et al. (2009) documented the prevalence of students who have difficulty writing in a population-based, birth cohort study. Specific writing disabilities were found to affect between 6.9% and 14.7% of students and existed with and without reading problems. Given this evidence that writing disabilities may affect approximately 10% of students, researchers and educators need more information to better understand early writing learning disabilities.

A more comprehensive understanding of writing disabilities is needed as expectations for student writing performance have increased due to the adoption of the Common Core State Standards (National Governors Association Center for Best Practices, 2010). The standards have specified end-of-the year expectations for producing different types of texts (narrative, opinion, and informative/explanatory texts), generating writing (responding to questions or texts), and using research to gather information. These standards are notable because they mark the first time that ambitious writing expectations have been adopted by a majority of US states (Shanahan, 2015). For students with writing disabilities, these standards present substantial challenges (Graham & Harris, 2013).

Students with Writing Disabilities

Learning to write, like learning to read, is complex and requires acquisition and integration of essential component skills. Writing researchers have demonstrated that several components are associated with writing development (see Berninger, 2009; McCutchen, 2006 for reviews). These include text generation (translating ideas into words, sentences, paragraphs, and discourse structures) and transcription (putting words, sentences, and higher levels of discourse into print). Below, we explain how these components relate to writing disabilities

Students with writing disabilities may experience difficulty with text generation, transcription or both. Dysgraphia is a writing-specific learning disability, and has been used more frequently to identify older elementary aged students. It is commonly characterized by inefficient and inaccurate transcription skills (e.g., handwriting and spelling) (Berninger, Nielsen, Abbott, Wijsman, & Raskind, 2008). However, difficulties with spelling are not unique to dysgraphia as they are also common for students with dyslexia (Berninger et al., 2008; Bishop & Snowling, 2004; Lyon, Shaywitz, & Shaywitz, 2003) and students with specific language impairments (Mackie, Dockrell, & Lindsay, 2013). Spelling can be difficult for students who may be at risk for writing problems. Spelling inefficiencies limit text production because students struggle to encode words. As a result, writing quality is compromised because of difficulties with handwriting and spelling (Abbott et al., 2010).

In addition to the challenges of transcription, students with writing problems also demonstrate difficulties associated with text generation. Older students with learning disabilities have been found to include fewer relevant ideas and more unrelated content in their writing (MacArthur & Graham, 1987). Students with language impairments may have smaller vocabularies, less mature syntactic complexity, and be less proficient at retrieving linguistic information necessary for text generation (Dockrell, Lindsay, Connelly, & Mackie; 2007; Mackie et al., 2013). Students with writing problems also devote less time to planning and revising their text than more proficient writers (Troia, 2006).

Understanding Beginning Writing Disabilities

Given the importance of writing and the number of students who could be at risk for writing disabilities, effective methods of identifying young students are needed. Researchers have faced substantial challenges in their efforts to identify struggling writers and to describe the characteristics or profiles of students with writing disabilities. A number of methodological approaches have been used. In some cases, researchers have relied on cut scores on a single assessment to identify students who might have writing difficulties (Costa, Hooper, McBee, Anderson, & Yerby, 2012; Hooper, et al., 2013). In other studies, researchers have used a discrepancy definition (Berninger et al., 2006; Berninger et al., 2008). Researchers have also considered more than one method for identifying struggling writers, including the use of a cut score and teacher ratings (Coker & Ritchey, 2014; Ritchey & Coker, 2014). A common limitation of these approaches is reliance on a single assessment to define either the nature of a writing disability or to identify a student as having a writing disability.

Other approaches that have been used to identify profiles of writing performance are multivariate methods such as cluster analysis (Roid, 1994; Hooper, Wakely, de Kruif, & Swartz, 2006; Wakely, Hooper, de Kruif, & Swartz, 2006). Roid (1994) identified eleven clusters in a large sample of third- and eighth-grade students who wrote an essay in one of five modes. All essays were scored with the same 5-point, analytic rubric. The clusters included groups with high and low performance across all six scores. Nine other clusters were identified, each with uneven performance across the analytic domains. Wakely and colleagues (2006) identified six clusters in a group of fourth- and fifth-grade students who wrote two narrative prompts and completed a reading assessment. The narratives were scored for quality, grammar, semantics, and spelling.

These clusters were characterized as average writers, expert writers, poor text quality, low spelling and reading, low grammar, and low semantics. Similarly, Hooper et al. (2006) found seven clusters using a range of assessments across four domains: problem solving, language, attention, and self-monitoring. Some clusters were characterized by average performance across the domains, but others indicated notable strengths or weakness in one or two domains (e.g., problem solving strength, problem solving weakness, and problem solving and language weakness). In these studies, clusters were estimated based on cognitive and linguistic measures (Hooper et al., 2006) or a combination of linguistic and reading measures (Wakely et al., 2006).

Purpose and Research Questions

We were interested in whether there were unique profiles of beginning writers based on norm-referenced assessments in writing that are often used to identify students with learning disabilities. This approach to identifying profiles of young writers builds on the use of multiple measures. We selected norm-referenced measures of writing related to transcription and text production. In addition, we wanted the measures to be sensitive to different levels of language because writing development has been shown to occur at various levels of language (Whitaker, Berninger, Johnston, & Swanson, 1994). The measures selected are reliable and valid assessments of spelling and sentence- and discourse-level writing. Previous work identifying writing profiles has been done with students in third grade and above (Hooper et al., 2006; Roid, 1994; Wakely et al., 2006), and we were interested in investigating the nature of writing profiles for beginning writers in first grade. With a better understanding of the profiles of young writers, it may be possible to identify struggling writers before their needs become difficult to remediate. Knowing more about the profiles of young writers may also help researchers identify targets for interventions.

In the current study, we used Latent Profile Analysis (LPA; Muthén, 2004) to investigate patterns in the writing performance of first-grade students. LPA is a person-centered analytic approach, and it focuses on relations among individuals in order to sort them into similar groups based on patterns occurring within the sample. While LPA is related to both factor analysis and cluster analysis, it has advantages over both approaches. LPA is a model-based approach that provides more flexible model specification. LPA classifies individuals based on the probability of group membership; in contrast, cluster analysis categorizes individuals using a dichotomous classification process (Pastor, Berron, Miller, & Davis, 2007). Finally, researchers can analyze the fit indexes of different models to select the most appropriate number of latent profiles.

Additionally, we were interested in investigating the external validity of the writing profiles to determine if these profiles were related to student performance on writing activities that are typical of classroom expectations (Fletcher, Francis, & Morris, 1988). This involved comparing the profiles to another assessment of writing performance. We wanted a robust measure of end-of-the-year writing, so students were asked to respond to writing tasks in two genres. Each sample was scored in multiple ways to avoid the limitations of a single score and to capture the multi-dimensional nature of first-grade writing (Kim, Al Otaiba, Folsom, Greulich, & Puranik, 2014). These individual scores were then combined into theoretically- and empirically-based factor scores using confirmatory factor analysis. The differences on the factor scores were then analyzed by writing profiles.

The specific research questions were as follows:

1. What are the latent profiles of beginning writers and how are at-risk writers characterized?
2. How does student performance on discourse-level writing tasks differ by latent profiles?

Based on previous research in this area, several hypotheses were made. First, we anticipated that students' performance on norm-referenced writing assessments would not be captured by a single profile and that multiple profiles would emerge. It seemed likely that at least three profiles would emerge characterizing students who demonstrated below average, average, and above average performance across all of the measures. In addition, we also hypothesized that additional profiles for students with needs in a specific area (e.g., weak spelling or writing fluency) would be found. Our third hypothesis is that students' latent profiles would be related to their performance on discourse writing tasks. For example, we expected that students in a profile characterized by above average writing would receive higher writing scores than students in profiles associated with below average performance.

Method

Participants

First-grade students (*N* = 391) in the Mid-Atlantic region of the United States participated in the study. The students were drawn from 50 classrooms in 13 schools in three school districts across two school years. These school districts serve between 10,000 and 17,400 students in urban and suburban neighborhoods. The schools varied in size; the number of first-grade classrooms in each school ranged from two to six. The research team worked with school district personnel to select schools that would yield a sample that was representative in terms of location (suburban and urban), student socio-economic status, and school size.

The participating students represent a range of ethnic backgrounds, language status, and disability status (see Table 1). We relied on information from the school districts about students' disability status. In total, 10.7 % of the participants received special education services. A range of disabilities was reported with the most frequent being a learning disability (4.1%) and a speech/language impairment (3.8%). Student-level, socio-economic status (SES) information was not provided by the districts, and only school-level SES was available. During the summer of 2013, the state department of education revised its method for calculating student SES. This policy change had a substantial impact on school-level SES statistics, even though the participating schools did not experience large demographic changes during this time period. To facilitate comparisons between the two years of data collection, we used SES information for the participating schools using data from the second academic year. On average, just over half of the students (54.9%) qualified for free or reduced-price meals with a range of 15.9% to 84.8%.

Classroom Context

The participating classrooms each had fewer than 22 students. Three classrooms used a co-teaching model with two teachers in the classroom. In four classrooms, the original teachers were replaced with long-term substitutes. The adopted reading curricula varied across the schools. Most commonly teachers utilized Houghton Mifflin Harcourt's *Journey's* (*n* = 32) (Baumann et al., 2011) or Pearson Scott Foresman's *Reading Street* (*n* = 5) (Afflerbach et al., 2011). Ten classrooms did not use a published reading curriculum and three classrooms used *Discover Intensive Phonics for Yourself* (Lockhard & Eversole, 2006). For writing instruction, 22 classrooms used a writing curriculum that was integrated within the reading curriculum. Five teachers also used an adaptable writing curriculum resource, *Explorations in Nonfiction Writing* (Stead & Hoyt, 2011). Almost half of the teachers did not use a standard writing curriculum (*n* = 23).

Measures

Spelling. Spelling ability was measured using the Spelling subtest from Woodcock Johnson Tests of Achievement, Third Edition (WJ-III; Woodcock, McGrew, & Mather, 2001, 2007). Students are asked to write the letters or words dictated by the examiner. The internal reliability estimate was reported as .92 for six-year-old students and .91 for seven-year-old students (McGrew, Schrank, & Woodcock, 2007). Inter-scorer agreement was 100%.

Writing Fluency. Sentence writing fluency was measured using the WJ-III Writing Fluency subtest (Woodcock et al., 2001, 2007). Students are asked to write as many simple sentences as possible using stimulus pictures and three related words within 7 min. The reliability estimates for seven-year-old students is reported to be .72 (McGrew et al., 2007). Inter-scorer agreement was 99.5%; any conflicts were resolved before analysis.

Writing Samples. Writing proficiency was assessed using the WJ-III Writing Samples subtest (Woodcock et al., 2001, 2007). Students are asked to respond to prompts that become increasingly difficult in terms of the length, vocabulary, grammar and the conceptual knowledge (Woodcock et al, 2001). The test has a reliability estimate of .89 six-year-old-students and .86 for seven-year-old students and a validity coefficient of .63 with WJ-III Spelling (McGrew et al., 2007). Inter-scorer agreement was 95%; any conflicts were resolved before analysis.

Writing Prompts. Students were asked to respond to two writing prompts designed to elicit two different genres—narrative and a descriptive. The narrative prompt was, “Think about one of your favorite activities. Write a story about a time that you had fun doing this activity.” The descriptive prompt was, “Think about a person you know well. It could be someone in your family or a friend. Describe that person and tell what he or she is like to someone who doesn’t know him or her.” The examiner gave each student a pencil and lined paper with the prompt at the top of the page. The examiner then read the prompt aloud, and students were allowed 20 minutes to complete the task. When students finished, the examiner directed the students to reread and check their work.

The scoring process for both writing prompts involved five coding methods designed to capture different multi-dimensional aspects of writing (Kim et al., 2014). Both texts were scored for length, quality, contextualized spelling, syntactic complexity and mechanics. Before scoring, students' narrative and descriptive texts were transcribed to reduce bias for poor handwriting. When scoring for quality, spelling mistakes were also corrected to reduced bias during the scoring process.

Length. The length of each text was calculated as the total number of correctly or incorrectly spelled words. Words were counted using a word count formula in Microsoft Excel. All texts were examined for random strings of letters or sequences of nonsense words (e.g., *qlArqrsuus* or *MeaMyBIDBeISesMocaCat*). These nonsense words were excluded from the total (14 from narrative texts and 45 from descriptive texts were excluded, less than .5% of the sample).

Spelling. Contextualized spelling was measured by identifying the percentage of correctly spelled words in the narrative and descriptive texts. Inter-scorer agreement was calculated using 20% of the texts for each genre. Inter-scorer agreement was 99.3% for narrative spelling and 98.9% for descriptive spelling.

Quality. The quality of the narrative and descriptive texts was measured by using a 6-point holistic rubric to provide a global rating of the text. The quality rubric was designed to be sensitive to three dimensions: (1) topic and detail, (2) organization and supporting details, and (3) word choice. The same rubric was used to score both narrative and descriptive texts. Inter-scored agreement (± one point) was calculated for 100% of the texts. Inter-scorer agreement was 96.2% for narrative quality and 96.8% for descriptive quality. The Spearman rho correlation between scorers was .88 for narrative quality and .87 for descriptive quality.

Syntactic Complexity. Two scores of syntactic complexity were calculated: mean length of T-unit (MLT) and clausal density. T-units were defined as a single main clause (independent clause) and any subordinate clauses or phrases associated with it (Hunt, 1965). MLT is the average total number of words per T-unit. Words in sentence fragments were not included in the MLT calculation. MLT was calculated for each text.

Clausal density was calculated as a ratio of the total number of clauses divided by the total number of T-units. A clause was defined as a group of words that contains a subject and a verb (Puranik, Lombardino, & Altmann, 2008). For example, *He went to the store because he needed bread* was counted as two clauses: (1) *He went to the store* as the independent clause and (2) *because he needed bread* as a dependent clause. Sentence fragments were removed before scoring clausal density. The total number of clauses in each text was calculated and then divided by the total number of T-units. Inter-scorer agreement was calculated on 20% of the data and was found to be 93% for narrative MLT and clausal density and 92% for descriptive MLT and clausal density.

Mechanics. Mechanics assessed the correct use of beginning capitalization and terminal punctuation in each T-unit. Unlike the syntactic measures, sentence fragments were included. A fragment was expected to begin with a capital letter and end with terminal punctuation (except titles, which required only capitalization). Incorrect capitalization in the middle of sentences was not scored for mechanics. In compound sentences, punctuation (i.e., a comma) at the end of the first T-unit was not required. Run-on sentences were fairly common, so a rule was adopted that a compound sentence with more than two independent clauses was not counted as correct. When multiple T-units were joined by coordinating conjunctions, an initial capital and terminal punctuation was required after every two T-units.

To control for the length of written texts, the percentage of correct initial capitals and terminal punctuation was computed for the total number of T-units. In total, four mechanics scores were calculated—percentage of correct capitalization in narrative text, percentage of correct capitalization in descriptive text, percentage of correct punctuation for narrative text, and percentage of correct punctuation for descriptive text. Inter-scorer agreement (based on 20% of the samples) for narrative text was 93.9% for capitalization and 93.4% for punctuation. Inter-scorer agreement for descriptive text capitalization was 96.2% and punctuation was 98.3%.

Procedures

The participating students were assessed by trained research assistants. The assessments were conducted outside of the students' classroom in a quiet space in the hallway or in an unused classroom. To reduce fatigue, the assessments were spread over several sessions two to three days apart. All assessments were administered from mid-April to the end of May. WJ-III Spelling and Writing Fluency were individually administered, and WJ-III Writing Samples and the narrative and descriptive writing prompts were administered in small groups of three or four students. The order of the narrative and descriptive prompts was counterbalanced to reduce testing effects, and at least one day was given in between the two prompts to reduce the effect of fatigue.

Analysis Procedures

Latent Profiles. Latent profile analysis (LPA) was conducted to classify students into discrete writing profiles using Mplus version 7.1 (Muthén & Muthén, 1998-2013). *W* scores from WJ-III Spelling, WJ-III Writing Fluency, and WJ-III Writing Samples were first converted to z scores. A 3-class model was specified and compared to models with 1, 2, 4, and 5 classes. First, models were examined regarding overall quality and convergence. Then, models were compared on entropy (Muthén, 2004), with values closer to 1.0 indicating better classification quality. Finally, models were compared using a set of fit indices: (a) Akaike information criterion (AIC; Akaike, 1974), with lower values indicating better fit; (b) Bayesian information criterion (BIC; McLachlan & Peel, 2000), with lower values indicating better fit; (c) Sample size adjusted Bayesian information criterion, with lower values indicating better fit; and (d) the Vuong-Lo-Mendell-Rubin adjusted likelihood ratio test (LRT-A; Lo, Mendell & Rubin, 2001) with significant *p* values suggesting the current model provides better fit compared to the model with one fewer class.

Factor Analysis. We conducted a confirmatory factor analysis (CFA) to generate factor scores based on narrative and descriptive writing using Mplus version 7.1 (Muthén & Muthén, 1998-2013). Given previous research on the dimensionality of first-grade writing, (Kim et al., 2014) we hypothesized that multiple factors would be present in our data. A preliminary analysis using exploratory factor analysis indicated that a four-factor model was optimal for these data; the 14 writing scores were reduced to a four-factor model.

We hypothesized a second-order factor model in which genre-specific writing dimensions comprise the first-order factors and general writing dimensions comprise the second-order factors. However, simultaneous model estimation requires three first-order factors per second-order factor while the current study focuses on only two writing genres, narrative and descriptive (Gerbing & Anderson, 1984). A comparable factor model may be constructed replacing genre-specific first-order factors with correlated measurement residuals. This requires the inclusion of correlated measurement residuals, which reflect common measurement error and shared variation associated with data collection methods (Fornell, 1983). For these data, we specified a model with correlated residuals.

We used four measures of fit to evaluate the four-factor model: chi-square ratio (χ^2/df), root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker-Lewis index (TLI). Use of chi-square as a measure of fit is increasingly biased toward statistical significance with large sample sizes (Dickey, 1996; Kline, 2005; Stevens, 1996). However, the chi-square ratio is less sensitive to sample size; a value less than 3 is indicative of acceptable model fit (Klein, 2005). The RMSEA is an absolute fit index; a value of 0 indicates exact fit, values below 0.05 indicate close fit, and values below 0.08 indicate reasonable fit (Browne & Cudek, 1993). The CFI and TLI are incremental fit indices that typically range between values of 0 and 1; values greater than 0.90 traditionally indicate good fit while more recent research suggests values greater than 0.95 are a preferable indicator of model fit (Bentler & Bonett, 1980; Hu & Bentler, 1999).

Differences in Writing Factor Scores by Latent Profiles. After developing latent profiles and creating factor scores for each student, we conducted a one-way between-groups multivariate analysis of variance (MANOVA) to examine differences among profiles on dependent variables (four factors). Before conducting the MANOVA, the homogeneity-of-slopes and homogeneity of variance assumptions were tested. All possible combinations of profiles and outcomes were tested to determine where the significant differences were located. Due to the violation of equal sample sizes for each group and equal variances among the groups assumptions, we used the Games-Howell test.

Results

Table 2 presents the descriptive statistics on all writing measures. Table 3 presents the correlations between writing measures and scores. Of note, the standard scores for this sample are slightly above the mean (Mean = 100, *SD* = 15).

Latent Profiles of Writers. The LPA tested profile solutions of one to six profiles. Two plausible models were identified by Entropy, AIC, BIC, Adj. BIC, and LRT test values (five and six solution). Entropy, Goodness-of-fit measures, and classification percentages are shown in Table 4. LRT values favored the five-profile solution because LRT was statistically significant, indicating that considering a fifth profile improved fit compared to the four-profile solution. The six-profile solution did not provide a better model than the five-profile solution. Information criteria measures, Entropy, LRT tests values, and classification percentages are shown in Table 4.

Figure 1 provides a display of the five profiles, and Table 5 provides model-based means for the profiles. The five-profile model separates students into the following profiles: At Risk (P1), Low Fluency (P2), Low Writing (P3), Average (P4), Above Average (P5), The At Risk profile (profile 1) included 8.5% of the sample (*n* = 33), and these students scored about 1.5 *SD* below average on WJ-III Spelling and WJ-III Writing Samples and over 2 *SD* below average on Writing Fluency. When we examined the composition of this profile, 40% of the students in this profile received special education services for a variety of classifications (developmental delay = 4, speech and language impairment = 1, specific learning disability = 7, other health impairment = 1). Considering this group's performance across the three writing measures, it is likely to represent students who are at risk for writing disabilities.

The Low Fluency profile (profile 2) was the smallest group with only 4.6% (*n* = 18) of the sample. This is one of two profiles with uneven performance across the three assessments. Students in this profile scored in the average range on WJ-III Spelling and WJ-III Writing Samples (less than 0.5 *SD* below the sample average) but over 1.5 *SD* below average the sample average on WJ-III Writing Fluency. In this profile, 19% of the students received special education services (speech and language impairment = 2, other health impairment = 1).

The Low Writing profile (profile 3) was the third largest group represented by 18.2% (n = 71) of the sample. Students in this profile scored about 0.6 *SD* below the sample average on WJ-III Spelling and about 0.3 *SD* below the sample average on WJ-III Writing Fluency, but more that 1.5 *SD* below the sample average on WJ-III Writing Samples. Twenty-one percent of students in this profile received special education services (developmental delay = 3, speech and language impairment = 2, specific learning disability = 8, other health impairment = 1).

The Average profile (profile 4) was the largest group in the sample with 38.6% of the students (n = 151) classified as belonging to this group. Across the three measures, students in this sample scored close to the sample mean. In this profile, four percent of students received special education services (speech and language impairment = 5, emotional disturbance = 1).

The Above Average profile (profile 5) represented 30.1% of the sample (n = 118) and was characterized by scores that were at least one standard deviation above the sample mean on all three measures. Four percent of the students classified in the Above Average profile were receiving special education services (speech and language impairment = 5).

Factor Analysis of Narrative and Descriptive Text Scores. Figure 2 depicts the four-factor model with standardized estimates. The four-factor model is specified with a simple structure in which each measure is modeled to load on one of four hypothesized factors: Quality/Length, Syntax, Spelling, and Mechanics. The four-factor model includes six correlated measurement residuals based on the theorized factor structure. The chi-square ratio of 2.39 indicates the four-factor model adequately fits the data (χ^2 = 155.0, df = 65, p < .001). The

RMSEA indicates the four-factor model is of reasonable fit (RMSEA = .060, p = .093). Furthermore, the CFI and TLI indicate the model is of good fit (CFI = .940, TLI = .906).

Differences in Writing Factor Scores by Latent Profiles. A one-way between-groups MANOVA was performed to investigate profile differences in factor scores (see Table 6 for correlations among factor scores). Four dependent variables were used: Quality/Length, Spelling, Mechanics, and Syntax. The independent variable was profile [At Risk (P1), Low Fluency (P2), Low Writing (P3), Average (P4), and Above Average (P5)]. Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicollinearity. Some violations were noted. The spelling factor did not approach normality. We found four cases considered multivariate outliers through the examination of Mahalanobis distances. These cases were not removed, and they were distributed across profiles. Box's Test of equality of covariance matrices was significant.

There was a statistically significant difference between profiles on the dependent variables (see Table 7). Using Pillai's trace, there was a significant effect of profiles on factors, V = .42, F (16, 1544) = 11.19, p = .001; partial eta squared = .104. When the dependent variables were considered separately, all four factors Quality/Length, Spelling, Mechanics, and Syntax reached statistically significance. For Quality/Length, F (4, 386) = 52.25, p < .001, partial eta squared = .351; for Spelling, F (4, 386) = 12.85, p < .001; partial eta squared = .118; for Mechanics, F (4, 386) = 42.11, p < .001; partial eta squared = .304; and for Syntax, F (4, 386) = 8.36, p < .001; partial eta squared = .080.

Finally, a series of post-hoc analyses (Games-Howell's test) were performed to examine mean difference comparisons across the five profiles and the four writing factor scores (see Table 8). The results revealed that all post-hoc mean comparisons were statistically significant

($p < .05$). For Quality/Length, the Above Average profile scored significantly higher than all the other profiles, the Average profile scored higher than the At Risk and Low Writing profiles, and the At Risk profile scored significantly lower than all other profiles. For Spelling, the Above Average profile scored significantly higher than all the other profiles and the Average profile scored higher than the At Risk profile. There were no differences among the lowest three profiles. For Mechanics, the Above Average profile scored significantly higher than all the other profiles and the Average profile scored higher than the At Risk and Low Writing profiles. There were no differences among the lowest three profiles. Finally, for Syntax, the Above Average profile scored significantly higher than all the other profiles except the Low Fluency profile. There were no differences among the lowest three profiles.

Discussion

In this study, we were interested in how best to characterize first-grade writers, with a focus on students who may be at risk for writing disabilities. We used latent profile analysis because of its methodological advantages, and to our knowledge, it has not been previously used to identify profiles of writers in first grade. Our hypotheses about the types of profiles that would emerge and the relationship between the profiles and performance on narrative and descriptive writing tasks were partially confirmed. Overall, we found that at the end of first grade there are measurable differences between students with below average writing skills and those with average and strong skills. This aligns with previous research that demonstrated variability in children's writing skills as early as kindergarten (Kim et al., 2011) and first grade (Hooper et al., 2011; Kim, Puranik, & Al Otaiba, 2015). The profiles that were found share similarities with previous work on classifying school-aged writers (Hooper et al., 2006; Roid, 1994; Wakely, et al., 2006); however, there are some notable differences that are discussed. Furthermore, several of the profiles were related to students' performance on extended writing tasks, which supports existing research on the contribution of spelling and sentence-writing skill to proficiency with extended discourse (Graham et al., 2002; Abbott, Berninger, & Fayol, 2010; Wagner et al., 2011).

Profiles of First Grade Writers

We had hypothesized that writing achievement would not be classified as a unitary construct and that at least three unique profiles were be identified. This hypothesis was confirmed as results of the latent profile analysis revealed five distinct profiles of first-grade writers (At Risk, Low Fluency, Low Writing, Average, and Above Average). Students' performance on extended writing tasks, capturing dimensions such as quality/ length, spelling, mechanics, and syntax, differed by profile.

Students in the At Risk profile demonstrated below average performance on all three assessments. These assessments require transcription skills at the word (WJ-III Spelling) and sentence level (WJ-III Writing Fluency), and they require text generation at the word and sentence level (WJ-III Writing Samples). Difficulty with both transcription and text generation at these two levels of language suggests that students in this profile have weaknesses in multiple areas important for writing. Furthermore, 40% of students in this profile had already been identified by their school for some type of special education service.

The global writing difficulties identified in the At Risk group have implications for both writing development and instruction. Students in the At Risk profile scored significantly below students in the Average and the Above Average profiles on all four dimensions of writing (operationalized by the factor scores). These first-grade students wrote narratives and descriptions that were rated as lower quality and were shorter than those produced by students in the other four profiles. In addition, the narratives and descriptions of At Risk students contained a higher percentage of both spelling and mechanical errors and had less sophisticated syntax than those produced by students in the Average and Above Average profiles. Overall, students in the At Risk profile were producing extended texts that differed markedly from those produced by students in the Average and Above Average profiles. Weak performance on these composing tasks is important because these are consequential in the classroom. The CCSS set end-of-year expectations for composing in several genres, including narratives and descriptions. Students who have difficulty with these tasks would be unlikely to meet the standards, which could have undesirable consequences for students and schools. In order to addresses students' needs, school should consider how to provide effective, early writing instruction.

A comprehensive approach to instruction might be most effective with writers in the At Risk profile. Across the three writing assessments, students in this profile scored below average on measures of transcription (spelling), sentence-level productivity (writing fluency), and text generation. Instruction that targets the skills and processes related to these components of the writing process might be the most beneficial for struggling students (Berninger, 2009). In addition, intensive intervention may be needed to help students with and at-risk for writing disabilities catch up with peers.

Two other profiles emerged that included below average performance on one writing task, and students in these profiles did not appear to have writing difficulties that were as pervasive as those in the At Risk group. Students in the Low Fluency profile had significantly lower scores on WJ-III Sentence Writing Fluency, and students in the Low Writing profile had significantly lower scores on WJ-III Writing Samples. The evidence suggested that the difficulties faced by students in these two profiles were related to more specific areas of difficulty. These targeted areas of weakness are likely to impact overall writing performance, but they may be less severe than the global difficulties found in the At Risk profiles.

With respect to narrative and descriptive writing, students in the Low Fluency and the Low Writing profiles performed in the average to below-average range on the extended writing tasks. On all four factor scores these two profiles were not significantly different from each other. In Quality/Length both of these profiles were higher than those of students in the At Risk profile but lower than those of students in the Above Average profile. On Spelling, Mechanics, and Syntax, students in the Low Writing and Low Fluency profiles scored lower than students in the Above Average profiles, but in these areas they were not writing better texts than students in the At Risk profile. In both Quality/Length and Spelling, students in the Low Writing profile scored significantly lower than students in the Average profile, but the Low Fluency profile was not significantly different from the Average profile. Overall, it appears that students in the Low Writing profile performed slightly weaker on the extended writing tasks than students in the Low Fluency profile. However, both groups were consistently stronger than the At Risk profile. As evidence, the percentage of students receiving special education services in both the Low Fluency and Low Writing profiles was approximately half the rate of the At Risk profile.

Research on writing subtypes with older students has demonstrated a number of profiles with mixed performance across measures (Hooper et al., 2006; Roid, 1994; Wakely et al., 2006). In this sample, 89 students (22.8% of the sample) were classified as part of such a profile. An implication of an uneven skill profile is that using a single measure of writing proficiency to identify students for a writing disability would be likely to miss students' strengths and needs that might emerge through the use of multiple measures. As schools work to identify students' writing needs, the assessment of multiple components of writing may provide useful information for instruction. Furthermore, these results indicate that assessments of transcription skills should be included in any battery used to assess students' writing.

For students in the Low Fluency and Low Writing profiles, instruction might be more effective if it targets areas of specific need. For example, students in the Low Fluency profile might benefit from focused instruction on sentence writing skills. Work by Hooper and colleagues (2006) has provided preliminary evidence that such aptitude by treatment interactions may be present in writing, and future research should investigate the extent of these interactions with young writers.

The Average and Above Average profiles were characterized by consistent performance equal to or above the sample mean. In both groups, the percentage of students that qualified for special education services was much lower than the other three profiles—about four percent. Nearly all of these students were receiving services for a speech and language impairment, which could include articulation, fluency, or voice difficulties that may not have a direct impact on writing achievement. The consistent performance of these two profiles across the three norm-referenced assessments and on the extended writing tasks suggests that a single writing assessment might be sufficient to gauge their performance level. Similarly, comprehensive instruction designed to strengthen all areas of proficiency may be most effective with these groups.

Differences Among Developing Writers

Within first grade there were notable differences in the performance of students in each profile, and the profiles were able to capture meaningful differences in students' extended discourse. It is important to consider this in light of the types of measures used to model the latent profiles. A limitation of the WJ III in assessing writing is that items that are typically administered to first-grade students do not include a writing task any longer than a sentence. Despite this limitation, the At Risk, Average, and Above Average latent profiles differentiate students in ways that are sensitive to students' skill with extended discourse. It may be that the assessments used to create the profiles tap important component skills that are related to success with extended discourse. These skills may include transcription skills related to spelling and sentence writing fluency, text generation processes and knowledge sources required in broad writing tasks, and the executive function skills that regulate the writing process. Certainly, theoretical accounts of early writing (Berninger, 2009) and empirical investigations of writing predictors (Hooper et al., 2011; Kim et al., 2011) have identified the importance of these components.

It should also be noted differences in the performance on the extended writing tasks were not found for all of the profiles. In particular, performance on the descriptive and narrative tasks did not consistently differentiate students in the Low Writing and Low Fluency profiles from those in all other profiles. However, the differentiation on the extended tasks that was found for these two profiles followed expectations. For example, the texts written by students in the Low Writing profile were significantly lower than the Average profile in Quality/Length and higher than those of the At Risk profile. These results are not surprising because the Low Writing profile identified students who struggled with comprehensive writing tasks, although not quite as much as students in the At Risk profile. Across the four factor scores the Low Fluency profile was not significantly lower than the Average profile. Since none of the writing factor scores assessed writing fluency directly, the failure to differentiate the Low Fluency and Average profiles may not be surprising. It may be that the generous time limit (20 mins) for the extended writing tasks made it easier for students in this profile to compensate somewhat for their difficulties in sentence writing fluency. However, their abilities to compensate may have been limited because the Low Fluency profile was significantly lower than the Above Average profile on every factor score except mechanics. Overall, the validity evidence for the Low Writing and Low Fluency profiles was not strong but the results aligned with our expectations about the tasks that were involved. Clearly more investigation of the validity of these and other early writing profiles is needed.

The profiles identified in our analysis were quite different from the clusters reported in previous studies (Hooper et al., 2006; Roid, 1994; Wakely et al., 2006). Of the five profiles identified in our analysis, three of them—the At Risk, Average and the Above Average profiles—were characterized by consistent performance across assessments. Two profiles, the Low Fluency and Low Writing profiles, scored lower on one subtest than the other two. In contrast, other researchers identified more clusters of writers—between six and 11. Some of the clusters showed a stable performance across assessments, but many that were identified revealed an uneven pattern across the assessments that were used.

Differences between our findings and those of other researchers may be explained several ways. First, different measures were used to assess various aspects of writing skills across the studies. We used three norm-referenced assessments of writing skill to form latent profiles. These were selected because of their wide availability and because these assessments and similar ones have been used to identify students with writing problems (Costa et al., 2012; Ritchey & Coker, 2014). In other studies, assessments were selected to assess cognitive domains (Hooper et al., 2006) and others were used to assess linguistic features of writing (Roid, 1994), or linguistic features of writing and reading performance (Wakely et al., 2006). It is likely that the use of different assessments would yield different clusters or profiles of student writing performance.

The age or developmental level of the writers would also be likely to contribute to differences across studies. Other researchers who employed cluster analysis used data from students in third grade and above. The first-grade writers in this study have developing transcription skills manifested by higher rates of spelling errors (Bahr et al., 2012) and less efficient handwriting (Graham et al., 1998). These transcription challenges would be likely to constrain text production, resulting in texts that are shorter and less sophisticated than those produced by older writers (McCutchen, 2006). Other components important for writing would also be expected to be less well developed, including knowledge about strategy use, discourse knowledge and background knowledge (McCutchen, 2011). All of these differences between younger and older writers would be likely to impact the nature of the profiles that were found.

Limitations and Directions for Future Research

One limitation of this study was that all students were in first grade. As a result, it was not possible to assess student writing growth. With a longitudinal sample, researchers could assess whether writing profiles remain stable over time or if they change as students' writing develops. Even if the profiles remain steady over time, the relationship between the profiles and other measures of writing performance may change such that the differences between the At Risk and Average writers may change over time.

This analysis of latent profiles could also be expanded by including additional measures of early writing skill. The number and type of profiles that were found depend on the measures that are included in the model. In this study, commonly used, norm-referenced assessments of writing were selected intentionally. However, these assessments are unable to capture all the skills and knowledge sources important for writing success.

In the future, additional writing assessments should be used to examine students' writing profiles. For example, better measures of the component skills of spelling, such as orthographic and morphological knowledge and phonological awareness, and components of handwriting, such as fine motor skills and attention, might be useful to explore. Some of these skills were used in previous cluster analyses with success (Wakely et al., 2006). Measures of writing motivation, self-efficacy and strategic knowledge would also provide important information that could contribute to a deeper understanding of the profiles of young writers. Currently there is no widely accepted, comprehensive measure of writing proficiency, and until researchers agree on a single measure or a battery of assessments, differences in students' skill profiles are likely to emerge.

In the future, research should seek to replicate these latent profiles with other samples. Furthermore, classifying students into profiles might be useful for intervention. Researchers could then investigate whether aptitude by treatment interactions are present with early writers in an effort to design tailored interventions.

Conclusion

In sum, students in first grade were characterized as fitting into one of five latent writing profiles, and there were differences in writing dimensions across most profiles. The findings signal the need for increased focus on writing development and instruction (including intensive intervention) in the early grades. If wide disparities among students' writing are present as early as first grade, it is likely that these differences will persist without effective intervention. Writing researchers should devote more attention to these issues, and teachers should work to implement effective instructional approaches in their classrooms.

References

Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology*, *102*(2), 281–298.

Afflerbach, P., Blachowicz, C., Boyd, C. D., Cheyney, W., Juel, C., Kame'enui, E., …Wixon, K. K. (2011). *Reading Street*. Glenview, IL: Pearson Scott Foresman.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions.* AC-19 (6): 716–723.

Bahr, R. H., Silliman, E. R., Berninger, V. W., & Dow, M. (2012). Linguistic pattern analysis of misspellings of typically developing writers in grades 1-9. *Journal of Speech, Language, and Hearing Research*, *55*, 1587–1599.

Baumann, J. F., Chard, D. J., Cooks, J., Cooper, J. D., Gersten, R., Lipson, M., … Vogt, M. (2011). *Journeys*. Orlando, FL: Houghton Mifflin Harcourt.

Berninger, V. W. (2009). Highlights of programmatic, interdisciplinary research on writing. *Learning Disabilities Research & Practice*, *24*(2), 69–80.

Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Writing problems in developmental dyslexia: Under-recognized and under-treated. *Journal of School Psychology*, *46*(1), 1–21. http://doi.org/10.1016/j.jsp.2006.11.008

Berninger, V. W., Rutberg, J. E., Abbott, R. D., Garcia, N., Anderson-Youngstrom, M., Brooks, A., & Fulton, C. (2006). Tier 1 and Tier 2 early intervention for handwriting and composing. *Journal of School Psychology*, *44*(1), 3–30.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588-606.

Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin, 130*, 858–86.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Beverly Hills, CA: Sage.

Coker, D. L., & Ritchey, K. D. (2014). Universal screening for writing risk in kindergarten. *Assessment for Effective Intervention*, 39(4), 245-256. doi:10.1177/1534508413502389

Costa, L.-J. C., Hooper, S. R., McBee, M., Anderson, K. L., & Yerby, D. C. (2012). The use of curriculum-based measures in young at-risk writers: Measuring change over time and potential moderators of change. *Exceptionality*, *20*(4), 199–217. http://doi.org/10.1080/09362835.2012.724623

Dickey, D. (1996). Testing the fit of our models of psychological dynamics using confirmatory methods: An introductory primer. In B. Thompson (Ed.), *Advances in social science methodology*, (Vol. 4, pp. 219-227). Greenwich, CT: JAI Press.

Dockrell, J. E., Lindsay G., Connelly V., & Mackie, C. (2007). Constraints in the production of written text in children with specific language impairments. *Exceptional Children*, *73*, 147–64.

Fletcher, J. M., Francis, D. J., & Morris, R. (1988). Methodological issues in neuropsychology: Classification, measurement, and non-equivalent group comparisons. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 1, pp. 83–110). Amsterdam: Elsevier Science.

Fornell, C. (1983). Issues in the application of covariance structure analysis: A comment. *Journal of Consumer Research*, *9*(4), 443-448.

Gerbing, D. W., & Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research*, *11*(1), 572-580.

Graham, S., Berninger, V., Weintraub, N., & Schafer, W. (1998). Development of handwriting speed and legibility in grades 1–9. *The Journal of Educational Research*, *92*(1), 42–52.

Graham, S., & Harris, K. R. (2013). Common Core State Standards, Writing, and Students with LD: Recommendations. *Learning Disabilities Research & Practice*, *28*(1), 28–37.

Graham, S., Harris, K. R., & Fink-Chorzempa, B. (2002). Contribution of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology*, *94*(4), 669–686.

Hooper, S., Costa, L.-J., McBee, M., Anderson, K., Yerby, D., Knuth, S., & Childress, A. (2011). Concurrent and longitudinal neuropsychological contributors to written language expression in first and second grade students. *Reading and Writing*, *24*(2), 221–252. http://doi.org/10.1007/s11145-010-9263-x

Hooper, S. R., Costa, L.-J. C., McBee, M., Anderson, K. L., Yerby, D. C., Childress, A., & Knuth, S. B. (2013). A written language intervention for at-risk second grade students: A randomized controlled trial of the process assessment of the learner lesson plans in a tier 2 response-to-intervention (RtI) model. *Annals of Dyslexia*, *63*(1), 44–64. http://doi.org/10.1007/s11881-011-0056-y

Hooper, S. R., Wakely, M. B., de Kruif, R. E. L., & Swartz, C. W. (2006). Aptitude-treatment interactions revisited: Effect of metacognitive intervention on subtypes of written expression in elementary school students. *Developmental Neuropsychology*, *29*(1), 217–241. http://doi.org/10.1207/s15326942dn2901_11

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

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

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78*, 243-255. doi:10.1037/0022-0663.78.4.243

Katusic, S. K., Colligan, R. C., Weaver, A. L., & Barbaresi, W. J. (2009). The forgotten learning disability: Epidemiology of written-language disorder in a population-based birth cohort (1976-1982), Rochester, Minnesota. *PEDIATRICS*, *123*(5), 1306–1313. http://doi.org/10.1542/peds.2008-2098

Kim, Y.-S., Al Otaiba, S, Folsom, J. S., Greulich, L., & Puranik, C. (2014). Evaluating the dimensionality of first-grade written composition. *Journal of Speech, Language, and Hearing Research*, *57*, 199–211. http://doi.org/10.1044/1092-4388 (2013/12-0152

Kim, Y.-S., Puranik, & Otaiba, S. A. (2015). Developmental trajectories of writing skills in first grade: Examining the effects of SES and language and/or speech impairments. *The Elementary School Journal*, *115*, 593–613. http://doi.org/10.1086/681971

Klein, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford.

Lo, Y., Mendell, N., & Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88, 767–778.

Lockhard, C., & Eversole, L. (2006). *Discover intensive phonics for yourself* (3rd ed.). North Salt Lake, UT: Reading Horizons.

Lyon, R. G., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, *53*, 1–14.

MacArthur, C. A., & Graham, S. (1987). Learning disabled students' composing under three methods of text production: Handwriting, word processing, and dictation. *The Journal of Special Education. 21*(3), 22–42.

Mackie, C. J., Dockrell, J., & Lindsay, G. (2013). An evaluation of the written texts of children with SLI: the contributions of oral language, reading and phonological short-term memory. *Reading and Writing*, *26*(6), 865–888. http://doi.org/10.1007/s11145-012-9396-1

McCutchen, D. (2006). Cognitive factors in the development of children's writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research.* (pp. 115–130). New York: Guilford Press.

McLachlan, G., & Peel, D. (2000). *Finite mixture models* (1st ed.). New York, NY: Wiley.

McGrew, K. S., Schrank, F. A., & Woodcock, R. W. (2007). *Technical manual*. *Woodcock-Johnson III Normative Update*. Rolling Meadows, IL: Riverside Publishing.

Muthén, B. O. (2004). Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences (pp. 345–368).* Newbury Park, CA: Sage.

Muthén, L. K., & Muthén, B. O. (1998–2013). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén and Muthén.

National Governors Association Center for Best Practices, Council of Chief State School Officers (2010). Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects K–5. National Governors Association Center for Best Practices, Council of Chief State School Officers, Washington D.C. Retrieved November 27 2012, from http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf

Office of Special Education Programs Report to Congress, 2014 accessed at http://www2.ed.gov/about/reports/annual/osep/2014/parts-b-c/36th-idea-arc.pdf

Pastor, D. A., Barron, K. E., Miller, B. J., & Davis, S. L. (2007). A latent profile analysis of college students' achievement goal orientation. *Contemporary Educational Psychology*, *32*, 8–47.

Puranik, C. S., Lombardino, L. J., & Altmann, L. J. P. (2008). Assessing the microstructure of written language using a retelling paradigm. *American Journal of Speech-Language Pathology, 17*, 107-120.

Ritchey, K. D., & Coker, D. L. (2014). Identifying writing difficulties in first grade: An investigation of writing and reading measures. *Learning Disabilities Research and Practice, 29*(2), 54-65. doi: 10.1111/ldrp.12030

Roid, G. H. (1994). Patterns of writing skills derived from cluster analysis of direct-writing assessments. *Applied Measurement in Education, 7*(2), 159-170.

Shanahan, T. (2015). Common Core State Standards: A new role for writing. *The Elementary School Journal*, *115*(4), 464–479. http://doi.org/10.1086/681130

Stead, T. & Hoyt, L. (2011). *Explorations in Nonfiction Writing: Grade 1*. Portsmouth, NH: Firsthand Heinemann.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Troia, G. A. (2006). Writing instruction for students with learning disabilities. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 324–336). New York: Guilford Press.

Wagner, R. K., Puranik, C. S., Foorman, B., Foster, E., Wilson, L. G., Tschinkel, E., & Kantor, P. T. (2011). Modeling the development of written language. *Reading and Writing*, *24*(2), 203–220. http://doi.org/10.1007/s11145-010-9266-7

Wakely, M. B., Hooper, S. R., de Kruif, R. E. L., & Swartz, C. (2006). Subtypes of written expression in elementary school children: A linguistic-based model. *Developmental Neuropsychology*, *29* (1), 125–159. http://doi.org/10.1207/s15326942dn2901_7

Whitaker, D., Berninger, V., Johnston, J., & Swanson, H. L. (1994). Intraindividual differences in levels of language in intermediate grade writers: Implications for the translating process. *Learning and Individual Differences*, *6*(1), 107–130. http://doi.org/10.1016/1041-6080(94)90016-7

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001, 2007). *Woodcock-Johnson III Tests of Achievement*. Rolling Meadows, IL: Riverside Publishing.

Table 1

Sample Demographics (N = 391)

		N	Percent
Gender	Female	203	51.9
	Male	188	48.1
Ethnicity			
	African American	112	28.6
	Asian	19	4.9
	Hispanic	48	12.3
	Native American	1	0.3
	White	198	50.6
	Other	13	3.3
ELL	Yes	34	8.7
	No	356	91
Special Education			
	Developmental Delay	7	1.8
	Emotional Disturbance	1	0.3
	Learning Disability	16	4.1
	Other Health Impairment	3	0.8
	Speech/Language Impairment	15	3.8
	None	349	89.3

Note. ELL = English Language Learner. Missing information for one student in the ELL category.

Table 2

Descriptive Statistics (N=391)

Measures	Mean	*SD*
WJ-III Spelling	108.69	13.33
WJ-III Writing Fluency	109.98	19.35
WJ-III Writing Samples	113.17	12.39
Descriptive Writing		
Length	36.12	26.42
Quality	3.25	0.88
% of Correctly Spelled Words	82.35	12.73
Mean Length of T-Units	6.16	2.22
Clausal Density	1.08	0.27
% of Correct Capitalization	58.38	32.92
% of Correct Punctuation	64.56	33.40
Narrative Writing		
Length	34.44	24.96
Quality	3.30	1.01
% of Correctly Spelled Words	80.05	14.49
Mean Length of T-Units	7.96	5.33
Clausal Density	1.13	0.41
% of Correct Capitalization	66.09	34.38
% of Correct Punctuation	68.07	33.77

Note. WJ-III = Woodcock Johnson Tests of Achievement, 3rd Edition. Scores for the WJ-III subtests are standard scores. Descriptive and Narrative length is number of words. Descriptive and Narrative quality is a 6-point holistic scale. Descriptive and Narrative % of correctly spelled words is the average percent of words spelled correctly in each text. Descriptive and Narrative mean length to T-units is number of words. Descriptive and Narrative clausal density is the average number of clauses per T-unit. Descriptive and Narrative % of correct capitalization is the average number of T-units with correct capitalization. Descriptive and Narrative % of correct punctuation is the average number of T-units with correct punctuation.

Table 3

Correlations among Norm-referenced Measures and Extended Writing Task Measures

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	WJ-III Spelling																
2	WJ-III Writing Fluency	.68															
3	WJ-III Writing Samples	.61	.58														
4	Descriptive Writing: Length	.16	.26	.17													
5	Descriptive Writing: Quality	.43	.49	.40	.63												
6	Descriptive Writing: % of Correctly Spelled Words	.59	.46	.44	.05	.29											
7	Descriptive Writing: Mean Length of T-units	.26	.14	.17	.18	.19	.13										
8	Descriptive Writing: Clausal Density	.20	.14	.20	.21	.26	.13	.67									
9	Descriptive Writing: % Correct Capitalization	.24	.23	.23	.05	.17	.14	.06	.03								
10	Descriptive Writing: % Correct Punctuation	.28	.24	.26	.05	.19	.22	.17	.14	.40							
11	Narrative Writing: Length	.33	.39	.31	.34	.42	.18	.10	.15	.18	.09						
12	Narrative Writing: Quality	.45	.46	.44	.31	.53	.28	.13	.18	.21	.17	.75					
13	Narrative Writing: % of Correctly Spelled Words	.56	.37	.35	.12	.28	.53	.11	.08	.14	.18	.22	.31				

Table continues

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
14	Narrative Writing: Mean Length of T-units	.11	.12	.11	.01	.03	.09	.07	.08	.02	.04	.13	.08	.13			
15	Narrative Writing: Clausal Density	.17	.14	.19	.06	.08	.14	.11	.12	.05	.08	.13	.18	.16	.38		
16	Narrative Writing: % Correct Capitalization	.08	.06	.09	-.04	.01	.11	-.02	.03	.13	.17	.01	.08	.25	.11	.28	
17	Narrative Writing: % Correct Punctuation	.17	.17	.19	-.01	.08	.18	-.07	-.01	.09	.30	.02	.11	.33	.24	.32	.44

Note. r values > .09, significant at $p < .05$; r values > .13, significant at $p < .01$. $N = 391$.

Table 4

Criteria for assessing fit for different number of latent profile solutions.

	1 Profile	*2 Profiles*	*3 Profiles*	*4 Profiles*	*5 Profiles*	*6 Profiles*
AIC	3337.83	3014.29	2881.95	2847.28	2810.57	2777.40
BIC	3361.64	3053.98	2937.51	2918.72	2897.88	2880.58
Adj. BIC	3342.60	3022.25	2893.09	2861.61	2828.08	2798.09
Entropy	*n/a*	0.769	0.804	0.793	0.825	0.860
LRT Test	*n/a*	331.53 *p* = .001	134.70 *p* = .020	40.95 *p* = .354	42.92 *p* = .008	39.52 *p* = .191
% for each profile	*P1* = 100%	*P1* = 38% *P2* = 62%	*P1* = 12% *P2* = 44% *P3* = 44%	*P1* = 10% *P2* = 42% *P3* = 16% *P4* = 32%	*P1* = 8% *P2* = 4% *P3* = 17% *P4* = 40% *P5* = 30%	*P1* = 8% *P2* = 17% *P3* = 4% *P4* = 34% *P5* = 33% *P6* = 4%

Note: *N*= 391. Model fit improves as AIC and BIC values decrease and entropy values approach one. Statistically-significant LRT indicate that the inclusion of an additional class improves model fit.

Table 5

Means and standard deviations for the five profiles

Scores	Five-Profile Solution				
	Below Average	*Low Writing Samples*	*Low Writing Fluency*	*Average*	*Above Average*
WJ-III Spelling	91.53	103.81	100.01	107.69	120.28
	(9.39)	(7.64)	(10.32)	(10.08)	(10.05)
WJ-III Writing Fluency	81.75	109.13	85.22	112.06	129.03
	(7.99)	(8.52)	(9.08)	(9.96)	(10.08)
WJ-III Writing Samples	89.00	94.38	109.51	114.62	122.40
	(10.45)	(6.57)	(7.86)	(8.45)	(6.48)

Note: Scores for the WJ-III subtests are standard scores.

Table 6

Pearson Correlations Associated with the Writing Factor Scores

	Quality /Length	Syntax	Spelling	Mechanics
Quality/Length	1.0			
Syntax	.457	1.0		
Spelling	.643	.514	1.0	
Mechanics	.270	.800	.589	1.0

Note: N=391

Table 7

Multivariate Between-Subjects Effects

Variables	*F*	*df*	Error *df*	Partial eta Squared	*p*
Quality/Length	52.246	4	386	.351	.000
Syntax	12.854	4	386	.118	.000
Spelling	42.112	4	386	.304	.000
Mechanics	8.361	4	386	.080	.000

Table 8

Factor Score Means, Standard Deviations, and post-Hoc Tests

	At Risk (1)	Low Fluency (2)	Low Writing (3)	Average (4)	Above Average (5)	Post-hoc
Quality /Length	-0.79 (0.52)	-0.23 (0.6)	-0.39 (0.58)	-0.01 (0.51)	0.47 (0.5)	1 < 2, 3, 4, 5 2 < 5 3 < 4, 5 4 < 5
Syntax	-0.37 (0.51)	-0.17 (0.44)	-0.15 (0.55)	0.02 (0.45)	0.18 (0.34)	1 < 4, 5 2 < 5 3 < 5 4 < 5
Spelling	-0.75 (0.83)	-0.21 (0.74)	-0.42 (0.82)	-0.03 (0.53)	0.51 (0.43)	1 < 4, 5 2 < 5 3 < 4, 5 4 < 5
Mechanics	-0.34 (0.56)	-0.1 (0.61)	-0.11 (0.56)	0.01 (0.45)	0.15 (0.38)	1 < 4, 5 - 3 < 5 4 < 5

Notes: N=391. The numbers in parenthesis in column heads refer to the numbers used for illustrating significant differences in the “Post-hoc” column.

Figure 1

Estimated score means: Five-Profile Solution

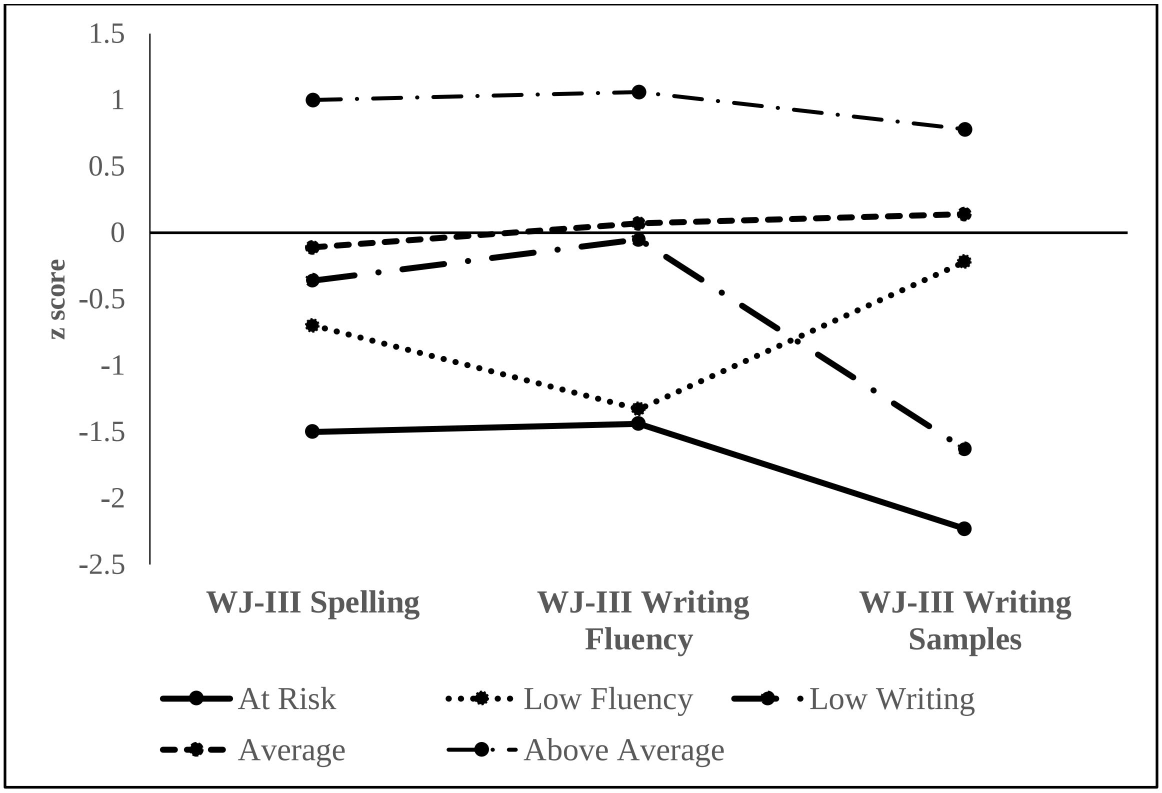
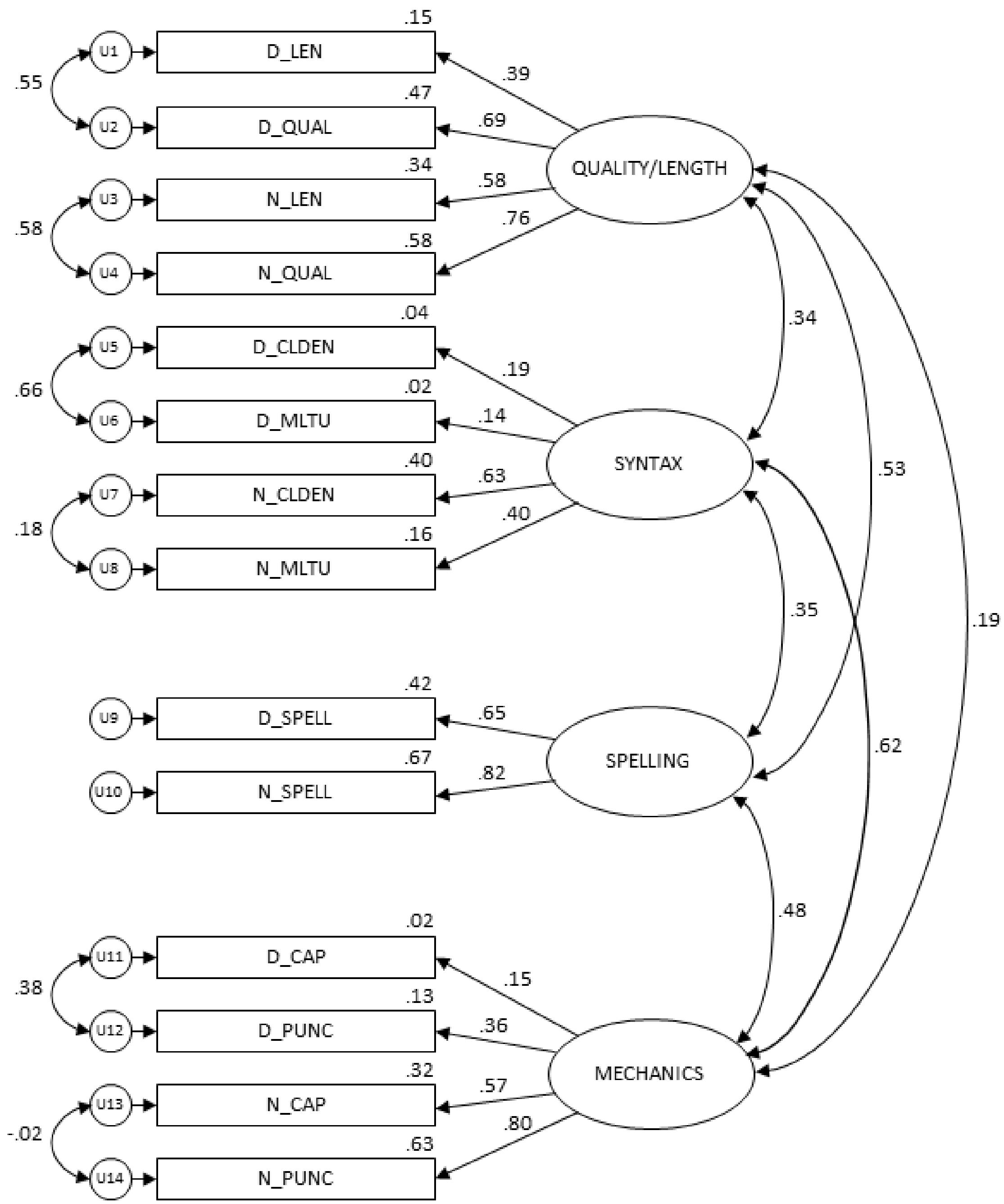


Figure 2

Four-Factor Writing Assessment Model with Standardized Estimates



Note: D_LEN = descriptive length; D_QUAL = descriptive quality score; N_LEN = narrative length; N_QUAL = narrative quality score; D_CLDEN = descriptive clausal density; D_MLTU = descriptive mean length of T-units; N_CLDEN = narrative clausal density; N_MLTU = narrative mean length of T-units; D_SPELL = descriptive percent correct spelling; N_SPELL = narrative percent correct spelling; D_CAP = descriptive percent correct capitalization; D_PUNC = descriptive percent correct punctuation; N_CAP = narrative percent correct capitalization; N_PUNC = narrative percent correct punctuation; U = unique variances.