Published in final edited form as:

J Res Educ Eff. 2017 ; 10(3): 619–645. doi:10.1080/19345747.2016.1237597.

Latent Profiles of Reading and Language and Their Association With Standardized Reading Outcomes in Kindergarten Through Tenth Grade

Barbara R. Foorman, Yaacov Petscher, Christopher Stanley, and Adrea Truckenmiller

Florida State University, Tallahassee, Florida, USA

Abstract

The objective of this study was to determine the latent profiles of reading and language skills that characterized 7,752 students in kindergarten through tenth grade and to relate the profiles to norm-referenced reading outcomes. Reading and language skills were assessed with a computer-adaptive assessment administered in the middle of the year and reading outcome measures were administered at the end of the year. Three measures of reading comprehension were administered in third through tenth grades to create a latent variable. Latent profile analysis (LPA) was conducted on the reading and language measures and related to reading outcomes in multiple regression analyses. Within-grade multiple regressions were subjected to a linear step-up correction to guard against false-discovery rate. LPA results revealed five to six profiles in the elementary grades and three in the secondary grades that were strongly related to standardized reading outcomes, with average absolute between-profile effect sizes ranging from 1.10 to 2.53. The profiles in the secondary grades followed a high, medium, and low pattern. Profiles in the elementary grades revealed more heterogeneity, suggestive of strategies for differentiating instruction.

Keywords

reading comprehension; reading profiles; academic language; latent profile analysis

The idea of targeting reading instruction to profiles of students' strengths and weaknesses in component skills is central to teaching. However, these profiles are often based on unreliable descriptions of students' oral reading errors (e.g., Argyle, 1989), text reading levels (Holliman, Hurry, & Bodman, 2014), or learning profiles (e.g., Tomlinson, 1999). Some reading researchers highlight components of word recognition and language skills within the Simple View of Reading as most representative of the teachable components of reading comprehension (Foorman, Herrera, Petscher, Mitchell, & Truckenmiller, 2015a; Foorman, Koon, Petscher, Mitchell, & Truckenmiller, 2015b; Wagner, Herrera, Spencer, & Quinn, 2014). To study profiles of these skills, researchers have used regression-based techniques to quantify the profiles of good and poor readers (Catts, Fey, Zhang, & Tomblin, 1999; Catts, Adlof, & Weismer, 2006) and profiles within poor readers (e.g., Badian, Duffy, Als, &

McAnulty, 1991; Buly & Valencia, 2002; Morris et al., 1998; Shankweiler et al., 1995; Stanovich & Siegel, 1994; Vellutino et al., 1996). However, these regression-based approaches typically use arbitrary achievement cut points, such as below the 30th or 40th percentile on a norm-referenced reading achievement test or an IQ-achievement discrepancy of 1.5 SDs, to define reader groups and, therefore, suffer from problems of reliability and generalizability. More recent approaches have taken a latent class approach (LCA) to modeling the observed measures to obtain reliable classes or profiles of reader characteristics. To date, the research employing LCA has focused on low-performing readers—those with language impairment (e.g., Catts, Compton, Tomblin, & Bridges, 2012; Justice et al., 2015), those in low-performing schools (e.g., Logan & Petscher, 2010), or struggling readers (Brasseur-Hock, Hock, Kieffer, Biancarosa, & Deshler, 2011). The current study contributes to research on deriving latent profiles of reading and language by having a representative sample of students in kindergarten through 10th grade and by relating the profiles to performance on important reading achievement outcomes.

Background on Profiles of Reading and Language Component Skills

Descriptive Approaches to Identifying Reader Profiles.

For decades practitioners have used the notion of learning profiles to describe patterns of oral reading errors (e.g., Argyle, 1989; Holliman et al., 2014) and performance on reading inventories (e.g., Denton, Ciancio, & Fletcher, 2006; McKenna & Stahl, 2009). Diagnostic inventories often describe skill profiles in terms of the five components of the National Reading Panel—phonemic awareness, phonics, fluency, vocabulary, and comprehension (National Institute of Child Health and Human Development, 2000). Skill inventories embedded in core reading programs typically reference state standards. With the recent adoption of new state standards in most states, teachers have to expand their conception of learning profiles to include notions of close reading, text difficulty, warrants and claims of disciplinary text, learning progressions, and academic language (Foorman & Wanzek, 2015; Heritage, 2010).

A framer of the practitioner discourse on “learning-profile differentiation” is Tomlinson (1999, p. 6). She and her colleagues recommend the “use of pre-assessment, self-assessment, and ongoing assessment to differentiate instruction for individual learning needs” (Brimijoin, Marquissee, & Tomlinson, 2003). She stresses that informative assessment is part of instruction and not just about readiness or outcome or finding weaknesses (Tomlinson, 2007). Although the identification of students who are struggling or at-risk has improved, matching intervention/instruction with specific data about instructional weaknesses is not widespread. The diverse needs at most schools require a range of interventions (Biancarosa & Snow, 2006). Many schools identify their poorest performing students and provide the same intervention program or class to all poor-performing students regardless of demonstrated variations in students’ needs. Building on Tomlinson’s recommendation for using assessment to inform instruction, researchers demonstrate that targeting instruction specifically to student profiles has positive impacts (Connor, Morrison, Fishman, Schatschneider, & Underwood, 2007). In order to study those profiles for instruction, researchers have used a variety of methods (e.g., regression, latent class

analysis) to identify those skills that are important to assess for informing instruction. These studies have been conducted in a variety of grade levels and identify the components of profiles by studying low-performing students or based on a theory of the components of reading.

Regression-based Approaches to Identifying Profiles of Poor Readers.

Reading psychologists have been studying the cognitive profiles of poor readers, often in pursuit of identifying causes of dyslexia but also to guide instruction. A popular approach was to compare students with and without IQ-achievement discrepancy on multivariate skills to highlight that phonological processing deficits were the common feature of dyslexia (e.g., Shankweiler et al., 1995). Using a regression-based, reading-level match design, Stanovich and Siegel (1994) found a phenotypic profile of reading-disabled children that they referenced in their phonological-core variable-difference model. According to this model, phonological coding was the cause of dyslexic students' poor word recognition, but there were other co-morbid cognitive difficulties, such as attentional problems. Morris et al. (1998) used cluster analysis to reliably identify seven subtypes of reading disability in 232 children. They found a core deficit in phonological awareness with discriminate validity on other measures involving phonological processing, language, and cognitive skills. Vellutino et al. (1996) showed that the cognitive profiles of easy-to-remediate poor readers differed from those of hard-to-remediate poor readers, thereby helping to make response-to-intervention (RTI) an alternative to the IQ-achievement discrepancy approach to identifying students as learning disabled (Vellutino, Scanlon, Small, & Fanuele, 2006).

Buly and Valencia (2002) applied both exploratory factor analysis and cluster analysis in deriving 10 profiles for their sample of 108 fourth-grade students who failed the Washington state reading test (WASL). The factor analysis was performed on a battery of reading measures and resulted in factors of word identification, meaning (comprehension and vocabulary), and fluency (rate and expression). The cluster analysis was performed on each student's three factor scores and other background information (i.e., WASL score and level, home language, poverty, writing ability). Eight of the resulting 10 clusters reflected low word identification and/or low fluency and the remaining two clusters consisted of students low on all three factors.

Each of these studies highlighted some malleable factors (e.g., phonological awareness, word recognition, vocabulary, fluency, comprehension) and some stable factors (e.g., attention, cognitive ability, poverty) that explain reading achievement for low-performing readers.

Latent Class Approaches to Identifying Reader Profiles.

The single-measure cut scores used in the regression-based studies described above allow for small changes in scores to result in shifting group membership and, hence, unstable classification. Francis et al. (2005) suggest a more robust approach of using more measures and maintaining the continuous nature of measures to infer membership in reading disability groups through latent classes. The use of multiple measures to form latent classes reduces measurement error and improves the reliability and stability of classification. A latent profile

approach (LPA) has benefits over the variable-centered approach of factor analysis, like those used in the regression-based approaches described previously, because LPA's person-centered approach models the probability of persons clustering together based on the similarity of their scores. LPA improves upon earlier regression-based approaches by using statistical tests and goodness of fit indices to determine student profiles (i.e., classes) instead of choosing an arbitrary cut score on assessments (Lubke & Muthén, 2005).

Several studies have used LPA to identify the reading and language profiles of students with impaired reading or language. Justice et al. (2015) examined the malleable emergent literacy skills of 218 preschool children with language impairment (LI). Four distinct profiles were identified, with the largest (55% of children) being low in all areas tested. The two low-risk classes consisted of children with higher language. The authors concluded that preschool children with poor language skills were at greater risk of future reading problems than children with high language skills.

In a study of a much later stage of reading development, Brasseur-Hock et al. (2011) used LPA with a sample of 319 students entering ninth grade to evaluate profiles of reading comprehension for adolescents. Students were administered three measures of reading comprehension, including the state test, and measures of timed and untimed decoding, oral reading accuracy and fluency, vocabulary, and listening comprehension. First they used LPA with the full sample of ninth grade students (ranging from proficient to struggling) to identify the students with poor reading comprehension. Next, they examined the skill profiles of students in the low-average and struggling-comprehender classes and identified five distinct skill profiles: Severe Global Weaknesses; Moderate Global Weaknesses; Dysfluent Readers; Weak Language Comprehenders; and Weak Reading Comprehenders. They point out that these five profiles of component skills demonstrate the importance of understanding the variability in component skills for targeting instruction and interventions with adolescents in addition to elementary-aged students. However, the development of reading skills throughout schooling indicates that better readers read more and through this greater exposure to print, vocabulary, syntax, and text structure, they gain better skills in each of those areas (Stanovich, 1986). Therefore, we expect to see a variety of profiles of strengths and weaknesses in the early grades and more consistent profiles of skills as students have more experience with schooling.

Are Language Skills Synonymous with Academic Language?

The researchers employing regression-based or latent class approaches to identify reader profiles tend to ground their work in the Simple View of Reading (Gough & Tunmer, 1986), whereby reading comprehension is the product of the unique contributions of word recognition and language. Simple View analyses from a latent variable perspective help clarify the dimensionality of language (Foorman et al., 2015a, 2015b; Language and Reading Research Consortium (LARCC), in press; Tomblin & Zhang, 2006) and the fading contribution of word recognition by grade 3 (Foorman et al., 2015a, 2015b). Components of language coalesce into one dimension in the primary grades (Mehta, Foorman, Branum-Martin, & Taylor, 2005; LARCC, in press) or a bi-factor model (Foorman et al., 2015b) above the primary grades.

How does this language dimension relate to academic language? Academic language refers to the language of schooling (Schleppegrell, 2012) and to a student's language proficiency in speaking, listening, reading, and writing (Uccelli, Galloway, Barr, Meneses, & Dobbs, 2015). Academic language is a register of language that differs from that of colloquial conversation. The clinical measures used by Simple View of Reading researchers are in the academic register and usually go beyond word recognition to capture the expanded construct of word knowledge important to academic language (e.g., knowledge of a word's pronunciation, spelling, multiple meanings, and structure). Clinical measures are less likely, however, to assess knowledge of the linguistic features that promote text cohesion (e.g., discourse connectors, anaphora, nominalization). Uccelli et al.'s (2015) Core Academic Language Skills (CALS) instrument as well as the assessment used in this investigation capture both the word and discourse knowledge components of academic language.

The Current Investigation

The previous studies summarized here highlighted specific skills and abilities that underlie important reading achievement outcomes. Those studies called for differentiated support and instruction for students in areas that would ameliorate their risk for future reading difficulties, particularly aligned to the Simple View of Reading. In order to hone the support for using assessment of component reading skills to inform instruction, the current study sought to contribute to this literature in several ways. First, this study utilizes a reading assessment battery that is efficient to administer but also provides reliable scores for low-stakes decisions when instructing component skills of reading. Second, the component reading skills assessment includes academic language, which has been shown to be an important predictor of future reading but is not consistently used. The assessment includes only skills that are malleable. In other words, educators are provided with a profile that facilitates choosing instructional materials and practices that directly impact students' reading skills. Although both cognitive abilities and motivation are important components of reading, providing formal assessment information about these skills is not necessary for teachers when choosing the content of their instruction. Finally, we demonstrate the malleable reading skills that are relevant for each grade level, kindergarten through tenth, and we incorporate all students in general education instead of focusing solely on struggling readers. The current investigation utilized latent profile analysis (LPA) to examine reading and language skills in a large, representative sample of Florida students in kindergarten through tenth grades. Additionally, it examined the relations among the latent profiles and gold standard reading achievement outcomes that are deemed important outcomes at each grade level.

Method

Participants

There were 7,752 participating students in kindergarten through tenth grades, 2,295 in kindergarten through second grade and 5,457 in third through tenth grades drawn from 41 schools in two large urban school districts in the Southeastern United States. The participating school districts administer a reading screening assessment and summative

reading assessments as part of their standard educational practices. Therefore, the sample of students represents all students participating in general education at the 41 schools. Numbers and demographics of participants by grade, gender, race/ethnicity, and English learner and Free-and-Reduced-Lunch status are provided in Table 1, as well as the reported number of students enrolled in kindergarten through grade 10 in Florida public schools and the state average demographics for comparison.

Measures

A component skills battery developed by the Florida Center for Reading Research (FCRR), called the FCRR Reading Assessment (FRA) was administered mid-year and a standardized reading test was administered at the end of the year during the 2012/13 school year (Foorman, Petscher, & Schatschneider, 2015c, 2015d). Additionally, scores on the state reading test were obtained for students in third through tenth grades. The FRA consists of a component battery for grades kindergarten through second grade and a component battery for third through tenth grades. In kindergarten through second grade, a flat, fixed-item, computer-administered version of the FRA was administered, with 30 items per task and a stop rule of four incorrect responses in a row. In third through tenth grades, a flat, fixed-item version of the reading comprehension task was administered but the other three tasks were administered in a computer-adaptive format.

FRA measures in kindergarten through second grade

Phonological Awareness—This kindergarten task requires students to listen to a word that has been broken into parts and then blend them back together to reproduce the word. Sample-based, marginal reliability (Lord & Novick, 1968) was estimated at .75. Concurrent validity is provided by a correlation of .36 with the Letter-Word Identification task of the Woodcock-Johnson III Test of Achievement (Woodcock, McGrew, & Mather, 2001).

Letter-Sounds—This kindergarten task requires students to provide the sound that a letter makes that is presented on the computer in upper and lower case letters. Sample-based, marginal reliability was estimated at .80. Concurrent validity is provided by a correlation of .52 with the Phonemic Awareness task of the Woodcock-Johnson III Test of Achievement (Woodcock et al., 2001).

Sentence Comprehension—This kindergarten task requires students to select the one picture out of four presented on the computer that depicts the sentence given by the computer (e.g., click on the picture of the bird flying towards the nest). Concurrent validity is provided by a correlation of .48 with the Sentence Structure subtest from the Clinical Evaluation of Language Fundamentals-4 (CELF-4; Semel, Wigg, & Secord, 2003). Sample-based, marginal reliability was estimated at .73.

Vocabulary Pairs—This task was administered in kindergarten through second grade. Three words (or pictures, for kindergarten students) appear on the monitor and are pronounced by the computer. The student selects the two words that go together best (e.g., *dark*, *night*, *swim*). Concurrent validity is provided by correlations with the Peabody Picture Vocabulary Test-4 (PPVT-4; Dunn & Dunn, 2007) of .46 in kindergarten, .59 in first grade,

and .50 in second grade. Sample-based, marginal reliability was estimated at .80 in kindergarten and first grade, and .70 in second grade.

Following Directions—This task was administered in kindergarten through second grade. Students listen, and click and drag objects in response to the computer's directions (e.g., put the square in front of the chair and then put the circle behind the chair). Concurrent validity is provided by correlations with the CELF-4 Concepts and Following Directions (Semel et al., 2003) of .58 in kindergarten, .58 in first grade, and .68 in second grade. Sample-based, marginal reliability was estimated at .83 in kindergarten, .80 in first grade, and .73 in second grade.

Word Reading—This task was administered in first and second grades. Words of varying difficulty (based on word frequency and item statistics) are presented on the monitor one at a time and students read them aloud. Sample-based, marginal reliability was estimated at .91 in first grade and .92 in second grade.

Spelling—This dictation task was administered in second grade. The computer provides a word and uses it in a sentence. Students respond by using the computer keyboard to spell the word. Words reflected second-grade spelling patterns contained in state curriculum standards, the scope and sequence of spelling programs, and research on spelling development (see Arndt & Foorman, 2010, for more information). The bivariate correlation between spelling and word reading was .78, similar to what has been found in other studies (Arndt & Foorman, 2010). Sample-based, marginal reliability was estimated at .91.

FRA measures in third through tenth grades.

Word Recognition—In this computer-adaptive task a word is pronounced by the computer and the student selects the word that best matches the pronunciation from a list of three words in a dropdown menu. Ten percent of the targets are nonwords in order to test grapheme-phoneme conversion skills (i.e., decoding) and the other 90% are real words selected based on grade-level frequencies (Zeno, Ivens, Millard, & Duvvuir, 1995). Distractors represent orthographic patterns known to predict reading comprehension (Garcia & Cain, 2014). For example, in the item with the target word “assembly,” the words in the dropdown menu were: *assembly*; *assemble*, *assembly*. Marginal reliability across third through tenth grades is .86, .88, and .93 for the fall, winter, and spring administrations and concurrent correlations range from .30 to .46 with Torgesen, Wagner, and Rashotte's (2012) *Test of Word Reading Efficiency* (Foorman et al., 2015d). In the present sample, 85.74% of students obtained an estimate of their word recognition ability with at least $\alpha = .90$ level of reliability.

Vocabulary Knowledge—In this computer-adaptive task a sentence appears on the computer screen with a word missing. The student selects from a dropdown menu the word that best completes the sentence. The words vary in morphological structure (i.e., affixes and roots). For example: *In some states you can get a driver's [permission, permissive, permit] when you are fourteen years old.* In this case the correct response (*permit*) is the root and the nonresponses are derivational suffixes. Evidence for this task's validity is provided by: 1) the

2–9% unique variance it accounted for in reading comprehension at the end of the year, controlling for prior reading comprehension, word analysis, and text reading efficiency (Foorman, Petscher, & Bishop, 2012); and 2) concurrent correlations with Dunn and Dunn's (2007) Peabody Picture Vocabulary Task-4th Edition ranging from .47 to .67 (Foorman et al., 2015d). Marginal reliability across third through tenth grades is .91, .89, and .90 for fall, winter, and spring performance. In the present sample, 99.42% of students obtained an estimate of their vocabulary knowledge ability with at least $\alpha = .90$ level of reliability.

Syntactic Knowledge—In this computer-adaptive task the computer reads aloud a sentence on the screen that has a missing verb, pronoun, or connective. The student selects from a dropdown menu the word from a list of three words in the same form class that best completes the sentence. An example of an item with connectives is: “Pizza is one of my favorite foods, [although, as, when] we only get to eat it on special occasions.” Verb tense, anaphora, and connectives are essential to text cohesion (Graesser, McNamara, Louwerse, & Cai, 2004). Evidence for the validity of connectives in predicting reading comprehension at the upper elementary and middle school levels has been demonstrated in several studies (e.g., Cain, Patson, & Andrews, 2005; Geva & Ryan, 1985; Uccelli et al., 2015; van Silfhout, Evers-Vermeul, Mak, & Sanders, 2014) and by and concurrent correlations ranging from .37 to .61 with the Grammaticality Judgment subtest of Carrow-Woolfolk's (2008) *Comprehensive Assessment of Spoken Language* (Foorman et al., 2015d). Marginal reliability across third through tenth grades is .93, .92, and .93 at the fall, winter, and spring administrations. In the present sample, 94.57% of students obtained an estimate of their syntactic knowledge ability with at least $\alpha = .90$ level of reliability.

Reading comprehension—Reading comprehension was measured by an experimental version of a computer-adaptive task called the Reading Comprehension Task. Students read four passages (both literary and informational) that varied in length from 200–1300 words and answered 7–9 multiple-choice questions per passage. Marginal reliability for the computer adaptive task across grades 3–10 is reported at .92, .88, and .93 for the fall, winter, and spring administrations, and demonstrates concurrent correlations ranging from .57 to .74 with the Stanford Achievement Test – 10th Edition (Foorman et al., 2015d). In the present sample, 86.60% of students obtained an estimate of their reading comprehension ability with at least $\alpha = .80$ level of reliability.

Standardized reading outcome measures—Kindergarten students were administered the Word Reading subtest of the Stanford Early Scholastic Achievement Test (SESAT), Form A, in May in small groups of five students. This subtest requires students to identify the printed name for a picture of an object after the name has been pronounced or identify a printed word after a spoken word has been pronounced. The Kuder-Richardson Formula 20 (KR20) reliability coefficient for spring administration is 0.85.

Students in grades 1–10 were administered the Reading Comprehension subtest of the Stanford Achievement Test—10th Edition (SAT-10; Harcourt Assessment, 2003) in their classroom in May. At the Primary 1 test level, first grade students (a) identify the picture described by a two-sentence story, (b) read short selections and demonstrate explicit and implicit understanding by completing sentences in a modified cloze format, and (c) read

short passages and answer multiple-choice questions tapping explicit and implicit information. The KR20 reliability coefficient for the Primary 1, form A, is 0.91. In grades 2–10, students take the Intermediate SAT-10 (grades 4–6) or the Advanced (grades 7–10). Students read literary, informational, and functional text passages and answer a total of 54 multiple choice questions that assess initial understanding, interpretation, critical analysis, and awareness and usage of reading strategies. The KR20 reliability coefficients for Form A range from .91-.92. Validity was established with other standardized assessments of reading comprehension, providing evidence of content, criterion-related, and construct validity (Harcourt Assessment, 2004).

In April, students in third through tenth grades took the state reading test, the Florida Comprehensive Assessment Test (FCAT 2.0). This group-administered, paper and pencil test consists of up to seven passages with multiple-choice questions aligned to four reporting categories: vocabulary, reading application, literacy analysis (fiction and nonfiction), and informational text/research process. Like the SAT-10, the FCAT 2.0 Reading questions require both literal and inferential understandings. Cronbach's alpha reliability coefficients range between .89 and .93 for the spring administration (Florida Department of Education, 2013). The publisher provides evidence of validity, including results of a second-order confirmatory factor analysis (CFA) that demonstrates that each grade-level assessment reflects a single underlying construct of reading achievement. Root mean square error of approximation (RMSEA) values range from .015 to .022, comparative fit index (CFI) values range from .98 to .99, and Tucker-Lewis index (TLI) values range from .98 to .99.

Procedure

Data were collected as part of a larger study including approximately 45,000 students in kindergarten through tenth grades during the winter and spring of 2013. In two 45-minute sessions over two days during the winter, students in kindergarten through second grades were individually administered the print-related tasks (Phonological Awareness and Letter Sounds in kindergarten, Word Reading in first and second grades, and Spelling in second grade) and the language tasks (Sentence Comprehension in kindergarten and Vocabulary Pairs and Following Directions in kindergarten through second grades). Classrooms of students in third through tenth grades were administered the computer-adaptive tasks—the Vocabulary Knowledge Task, the Word Recognition Task, and the Syntactic Knowledge Task—and the flat version of the Reading Comprehension Task in two 45-minute sessions over two days during the winter. During April, the SESAT Word Reading was administered by research staff to small groups of kindergarten students and the SAT-10 was administered by classroom teachers to all students in first grade through tenth grade as part of district assessment practices. Also during April, classroom teachers administered the reading portion of the FCAT over two consecutive days as part of state accountability testing.

A common-item non-equivalent groups design was used to collect data on the K-2 tasks and the baseline Reading Comprehension Task with 20% common items across forms. A planned missing data design was implemented such that all students were administered baseline Reading Comprehension, and differentially assigned to be administered the computer-adaptive version of Vocabulary Knowledge, Word Recognition, and Syntactic

Knowledge Tasks. The resulting ability score from an item response theory analysis was used for all four tasks. Raw scores from the FRA tasks in kindergarten through second grade were converted to z-scores for the purpose of analysis. The z-scores reflect the Florida norms. A latent factor score for reading comprehension was created from developmental scale scores from the RCT, SAT-10, and FCAT.

Data Analysis

A two-step process of latent profile analysis and general linear modeling were conducted at each grade level. A latent profile analysis (LPA) is a special case of mixture modeling that is conceptually similar to confirmatory factor analysis (CFA) and cluster analysis. The goal of LPA is to classify individuals in the sample into mutually exclusives classes, or profiles, based on their responses to the observed variables. Like a CFA, the LPA uses maximum likelihood estimation to estimate a latent factor that is assumed to cause the observed measures (i.e., the K-10 FRA tasks). The fit of the hypothesized model may be evaluated against other specified models using indices including the log likelihood, the Akaike Information Criteria (AIC), and the sample-size adjusted Bayesian Information Criteria (BIC). Where LPA differs from CFA is in the design of the latent variable (Samuelsen & Raczynski, 2013). When manifest variables are categorical or continuous but the latent factor is continuous, then the latent variable method falls into realm of CFA. However, if the latent factor is categorical, then the analysis becomes a mixture model; when the manifest variables are categorical it is referred to as latent class analysis, but with continuous variables, the model is referred to as LPA.

Similar to a cluster analysis, an LPA seeks to operationally separate individuals into classes, or profiles, based on the manifest variables. The primary difference between the two approaches is that the former seeks to generate clusters by evaluating how differences in mean performances are minimized within-cluster and maximized between-cluster. The resulting classification of individuals into clusters is based on how close the performance of an individual minimizes within-cluster differences. Conversely, LPA classifies individuals based on factor analysis; that is, LPA assumes that there is a relation among the manifest variables but that relation is unknown and caused by a categorical latent factor. LPA uses posterior probabilities of classification to describe the relation between individual and n profiles, such that final group classification is based on the class with the highest posterior probability of membership. LPAs were run at each grade level testing the fit of 2, 3, 4, 5, and 6 profile solutions. Each model was compared to others based on the AIC and BIC values, a log-likelihood ratio test, and substantively meaningful profile results.

Following the latent profile analysis, multiple regression analyses by grade level tested the extent to which profiles were statistically and meaningfully separated on the standardized measure of word reading in kindergarten and reading comprehension in all other grades. As a step to safeguard against bias in profile membership, the posterior probabilities of class membership were evaluated for magnitude. Because LPA assigns class membership based on the absolute magnitude of the poster probability, it is possible that an individual could be classified into a profile where their magnitudes across classes were really equal. For example, in a 3-class solution, an individual could have a .33, .33, and .34 posterior

probabilities of membership in classes 1, 2, and 3 and be assigned class 3 membership due to absolute magnitude. Given that there are no guidelines for acceptable posterior fit probabilities, in the present design a threshold of .70 was set so that relative confidence could be assured in testing for profile differences in the standardized outcome. Within-grade multiple regressions were subject to a linear step-up correction to guard against the false-discovery rate (Benjamini & Hochberg, 1995).

Results

Results are presented separately for the lower elementary grades (kindergarten through second) and upper elementary grades (third through fifth) because of the different nature of the assessments. Results from only one of the secondary grades—eighth grade—will be presented due to the similarity of results across these middle and high school grades. Results from the rest of the secondary grades are provided in supplemental materials.

Descriptive Statistics and Correlations

Kindergarten through second grade—Descriptive statistics and correlations among raw scores for the FRA measures and SESAT Word Reading and SAT-10 Reading Comprehension developmental scale scores are provided for kindergarten through second grade in Table 2. Not surprisingly, print-related measures were moderately correlated in kindergarten (Phonological Awareness with Letter Sounds, .48, and with SESAT Word Reading, .58; and SESAT Word Reading with Letter Sounds, .51). In first and second grades, print-related measures were more strongly correlated: SAT-10 with Word Reading (.75 in first and .58 in second grades); Word Reading and Spelling in second grade (.77). Academic language measures were moderately correlated with each other in all three grades and Vocabulary Pairs was moderately correlated with SAT-10 in first (.58) and second (.62) grades.

Third through fifth grades and eighth grade—Descriptive statistics and correlations for FRA measures and reading comprehension measures of RCT, SAT-10, and FCAT for third through fifth grades and eighth grade are presented in Table 3. Not surprisingly, the three reading comprehension measures were strongly correlated, with the RCT bivariate correlations ranging from .67 in eighth grade to .81 in fifth grade. FCAT and SAT-10 correlations ranged from a low of .71 in eighth grade to a high of .81 in third and fifth grades. VKT and SKT were moderately correlated in these grades (.31 to .46) as were the bivariate correlations of WRT (.29 to .51).

Latent Profile Analyses and Associations with Reading Outcome

The first research question explores the profiles that characterized students' literacy performance and the second research question demonstrates how the profiles were associated with the reading comprehension outcome. Fit indices from model testing are presented in Table 4. In all modeling cases, final model selection was based on both the model fit via AIC, BIC and the log likelihood test across models, as well as substantive theory. Optimal solutions in a few grades were characterized by a low within-profile π ; however, these were retained for the purposes of general linear model testing. Results from

Table 4 demonstrate that across grades a consistent, significant reduction in the log likelihood was observed when testing the difference between n and $n-1$ profiles. Along with this reduction were general reductions in the AIC and BIC values. In many instances, a class with a statistically significant reduction in the log likelihood was not selected for final class retention. This decision was due to the nature of latent profile analysis being an exploratory profile analysis. As such, certain profiles may have yielded better model fit, yet the class itself was either relatively homogeneous to other classes or had a low within-profile n . For example, in first grade, a six-profile solution fit significantly better than a five-class solution, but the additional class only contained three individuals and was not deemed a theoretically meaningful class. Likewise, model fit for eighth grade suggested that four to six profiles fit significantly better than a three-profile solution. As the results will detail, the profiles in later grades were not heterogeneous across the manifest variables and were only separated on the level of performance (i.e., high ability, average ability, or low ability across measures). Added profiles above the three-profile solution contributed little unique information as they simply reflected shades of high, average, or low ability.

Kindergarten through second grade.—Results of the latent profile analysis (LPA) for kindergarten through second grade are presented in Figure 1. Z -scores of FRA measures are presented on the Y-axis. General linear model comparisons are provided in Table 5, with the Critical p -value from the linear step-up correction noted in a separate column. There was one p -value in kindergarten ($p=.012$) that was not less than the critical p of .006 and, therefore was marked with an asterisk as non-significant. All remaining p -values were statistically significant after applying linear step-up correction.

Kindergarten: FRA measures in kindergarten consisted of three academic language tasks—Vocabulary Pairs, Following Directions, and Sentence Comprehension—and the print-related tasks of Phonological Awareness and Letter Sounds. There were six identified literacy profiles for the 422 kindergarten students (see top left of Figure 1) and they are described in the top of Table 6. Class 6, with 19% of students, and Class 3, with 42% of students, were above average on all variables, but Class 3 was 1 SD lower on Vocabulary Pairs. The other four classes tended to have low performance but varied on the relative strengths and weaknesses of language variables versus the print-related variables of Phonological Awareness and Letter Sounds.

These classes were distinctive both statistically and practically based on SESAT Word Reading performance, $F(5, 315)=37.11, p<.001$, accounting for 37% of variance. Not surprisingly, class 1 had significantly worse Word Reading outcomes than all other classes (performing below the 25th percentile) except class 5, and class 6 had significantly better Word Reading than all other classes (performing at the 75th percentile). Interestingly, outcomes did not differ between class 4 (average academic language) and class 5 (average print-related skills), nor did they differ between class 2, with high Phonological Awareness and low Sentence Comprehension, and classes 3,4, and 5. These classes tended to be below average on FRA literacy skills and below average on SESAT Word Reading. The average absolute value of the standardized difference in SESAT performance across all classes was

Hedge's $g = 1.10$ (Table 5), further reinforcing the magnitude of differences in component skill profile performance on the standardized outcome.

First grade: FRA measures in first grade consisted of Word Reading and the two academic language measures of Vocabulary Pairs and Following Directions. There were five identified literacy profiles for the 989 students in grade 1 (see top right of Figure 1) and the five profiles are described in the middle of Table 6. Class 2 (35% of students) was high, Class 5 (43% of students) was average, and Class 4 (17% of students) was low on all variables. Classes 1 and 3 were small in size and differed by more than 3 SD on Following Directions.

These classes were distinctive in SAT-10 Reading Comprehension, $F(4, 984) = 127.63$, $p < .001$, accounting for 34% of variance. Class 2, which had an above average profile on FRA tasks, performed above the 75th percentile on SAT-10, significantly better than all other classes. Class 5, with average FRA performance, had average performance on the SAT-10, which was significantly higher than all other classes except class 2. Class 4, which performed about 1 SD below average on FRA tasks, performed at about the 30th percentile on the SAT-10, significantly better than only one class—class 1. Classes 3 and 4 did not differ significantly in SAT-10 performance, both being at the 25th percentile, on average. Class 1, with its very low Following Directions score, performed below the 25th percentile on the SAT-10, significantly lower than all other classes. The average absolute value of the standardized difference in SAT-10 performance across all classes was Hedge's $g = 1.43$ (Table 5).

Second grade: FRA measures in second grade were the same as those in first grade (Vocabulary Pairs, Following Directions, and Word Reading), but also included Spelling. There were 884 participating students in second grade and six profiles were identified. As can be seen from the description of the patterns in the bottom of Table 6 and the bottom of Figure 1, Class 5 (32% of students) was high on all measures, Class 3 (15% of students) was low, and Class 1 (5% of students) was very low. Class 6 (5% of students) was very low on the print-related variables. Classes 2 (10% of students) and 4 (32%) were mirror images of each other, with Class 2 high on print-related variables and low on language variables and Class 4 the opposite.

These classes were distinctive in SAT-10 Reading Comprehension, $F(5, 878) = 147.39$, $p < .001$, accounting for 46% of variance. Class 5, which had an above-average profile on FRA tasks, performed above the 75th percentile on SAT-10, significantly better than all other classes. Class 1, which had a very low-performing FRA profile, performed below the 25th percentile on SAT-10, significantly worse than all other classes. The classes with mirror image profiles that hovered around the FRA mean—classes 2 and 4—also performed at the 50th percentile of the SAT-10 and, therefore, did not significantly differ from each other. The relatively high academic language performance of class 6 did not compensate for its very low print-related performance on FRA tasks, nor for its low SAT-10 performance at the 25th percentile. Class 3 also scored at the 25th percentile of the SAT-10 and, therefore, was not significantly different from class 6 in this regard but was significantly lower than classes 2 and 4. The average absolute value of the standardized difference in SAT-10 performance across all classes was Hedge's $g = 1.48$ (Table 5).

Third through fifth grades and eighth grade.—Results of the latent profile analysis (LPA) for third through fifth and eighth grades are presented in Figure 2, depicting the z -scores of the FRA measures for the Word Recognition Task, the Vocabulary Knowledge Task, and the Syntactic Knowledge Task. The FRA's Reading Comprehension Task was not included in the LPA in order to avoid autoregressive effects. However, it was utilized as part of the latent outcome measure of reading comprehension, along with the SAT-10 and FCAT. General linear model comparisons among the latent profile classes and the corresponding effect sizes reflecting differences in the latent measure of reading comprehension are shown in Table 7. All p -values $< .05$ were statistically significant after applying linear step-up correction.

Third grade: Five literacy profiles were identified for the 607 third grade students (see top left of Figure 2). Descriptions of the patterns are provided in the top of Table 8. The vast majority of students (78%) were in class 3, performing at the mean on all three FRA tasks. Students in classes 4 and 5 (representing 8% and 5% of students, respectively) performed above average, but differed on the Word Recognition Task. Classes 1 and 2 (representing 5% and 4% of students, respectively) performed below average and were similar on Syntactic Knowledge but mirror images of each other on Vocabulary Knowledge and Word Recognition.

These classes were distinctive on latent reading comprehension, $F(4, 559)=44.30, p < .001$, accounting for 24% of the variance. The two classes that were above average on FRA measures, class 4 and class 5, both performed about 1 SD above the mean on reading comprehension and were not significantly different from each other but were significantly better than the other three classes. Class 3's and class 2's means were about 0.5 SD below average and were not significantly different from each other. Class 2 marginally outperformed class 1 (Hedges $g = -.53$). The average absolute value of the standardized difference in latent reading comprehension performance across all classes was Hedge's $g = 1.42$ (Table 6).

Fourth Grade: Five literacy profiles were identified for the 584 students in fourth grade (see top right of Figure 2). The descriptions of the patterns in Table 8 show that of the two above-average classes and the three below-average classes, differences were greatest on the Vocabulary Knowledge Task.

These five classes were distinctive on latent reading comprehension, $F(4, 327)=54.05, p < .001$, accounting for 40% of the variance. All classes were significantly different from each other except for the three classes performing below average—classes 1, 3, and 5—which all performed about 1 SD below average reading comprehension in spite of large differences in Vocabulary Knowledge. In contrast, the large difference in Vocabulary Knowledge between class 4 and class 2—almost 2 SD—may help explain why class 4 scored about 1.25 SD above the mean on reading comprehension, whereas class 2 performed only 0.5 SD above the mean. The average absolute value of the standardized difference in latent reading comprehension performance across all classes was Hedge's $g = 1.34$ (Table 6).

Fifth grade: Five literacy profiles were identified for the 659 students in fifth grade (see bottom right of Figure 2). The patterns described in Table 8 show that the vast majority of students' performance on all variables was low (Class 3 with 53% of students) or average (Class 4 with 35% of students). The 7% of students in Class 5 performed very high on all variables. Classes 1 and 2 performed low, but the 1% of students in Class 1 performed very low on Vocabulary Knowledge.

These five classes were distinctive on latent reading comprehension, $F(4, 482)=132.73, p < .001$, accounting for 52% of the variance. All classes were significantly different from each other, with their means on latent reading comprehension corresponding closely with their FRA means: class 5 and class 4 were above the mean, about 1.75 SD and about 0.5 SD, respectively; class 3 was about 0.5 SD, class 2 about 1.25 SD, and class 3 about 1.65 SD below the mean. The average absolute value of the standardized difference in latent reading comprehension performance across all classes was Hedge's $g = 2.53$ (Table 6).

Eighth grade: Three literacy profiles were identified in eighth grade (see bottom left of Figure 2 and descriptions in Table 8). These profiles fell into high, average, and low FRA profiles with corresponding high, average and low reading comprehension, and that was typical of the patterns in sixth through tenth grades, except for ninth grade where there were two above average and two below average, flat literacy profiles with corresponding reading comprehension. The amount of variance in reading comprehension accounted for by the FRA literacy profiles in these middle and high school grades ranged from a low of 41% in seventh grade to a high of 61% in ninth grade.

Out of the 629 students in eighth grade, Class 1 had the vast majority of students (72%) and they performed just below the FRA mean on all variables. Class 2, with 25% of students, performed about 1 SD above the mean on all variables. Class 3, with 3% of students, performed about 1.5 SD below the mean on Word Recognition and Syntactic Knowledge, but 2.5 SD below the mean on Vocabulary Knowledge.

These three classes were distinctive on latent reading comprehension, $F(2, 521)=196.02, p < .001$, accounting for 43% of the variance. All contrasts were significant. Reading comprehension means corresponded with FRA means: class 2 was about 1 SD above the mean; class 1 about 0.4 SD below the mean; and class 3 about 1.5 SD below the mean. The average absolute value of the standardized difference in latent reading comprehension performance across all classes was Hedge's $g = 2.19$ (Table 6).

Discussion

The current investigation had two objectives. The first objective was to determine the latent profiles of reading and language skills that characterized students in kindergarten through tenth grade. The second objective was to determine the extent to which these latent profiles were related to important reading outcomes in kindergarten through tenth grades.

Latent Profiles and Their Relation to Reading Outcomes Across the Grades

Latent profile analysis (LPA) identified five to six classes in the elementary grades and only three in the secondary grades. In all grades the latent profiles were significantly related to the norm-referenced reading outcome scores, accounting for a low of 24% of the variance in third grade to a high of 61% of the variance in ninth grade, with the mode being 42%. Although this level of variance is not as high as other studies that have used more time-consuming clinical measures (Buly & Valencia, 2002; Catts et al., 2012), the information obtained from the profiles in this study is substantively important given the effect size of these profiles. The range of average absolute values of the standardized difference in reading outcome across all latent classes in a grade using Hedges g was 1.10 in kindergarten to 2.53 in fifth grade. Previously, the field was dominated by latent class analyses of clinical samples (e.g., Catts et al., 2012; Justice et al., 2015) or low-performing students (Logan & Petscher, 2010; Brasseur-Hock et al. 2011). Because this is the first study that examined latent profiles of students covering the full range of reading performance (not just low-performing students) and spanning 11 grades, it is possible that other non-malleable components of reading comprehension account for more variance with higher-performing students.

The finding that the profiles above fifth grade fell into a pattern of low, medium, and high performance is noteworthy. In eighth grade, 25% of the sample of 629 students performed 1 SD above the mean on FRA measures and on latent reading comprehension, 72% performed .25 SD below the mean, and 3% performed 1.5 to 2.5 SDs below the mean. This very low group was relatively better on Word Recognition than on the academic language tasks (Syntactic Knowledge and Vocabulary Knowledge). By contrast, Brasseur-Hock et al. (2011) found that dysfluent readers and readers with severe or moderate global weaknesses were more common profiles than the profile of weak language comprehenders among their ninth grade students scoring at the lowest level of the Kansas state reading test. Their vocabulary measure, however, tapped breadth over depth, and they did not measure the syntactic knowledge inherent in text discourse as the FRA does. The eighth grade results presented here suggest that students with very low reading comprehension need intervention in word recognition but also need intervention on the academic language skills of vocabulary and text discourse. Other evidence suggests that these core academic language skills can and should be taught (Uccelli et al., 2015).

Instructional Implications of Latent Profiles in the Elementary Grades

The five to six reading and language profiles found in the elementary grades reflect heterogeneity of skills that should be taken into account when differentiating instruction. In kindergarten through second grade there was a small class with 1–7% of students who were low on all FRA measures and below the 25th percentile on reading outcome. This profile of global weakness suggests intervention in all alphabetic and academic language skills. But even among the low-performing classes there was sometimes a notable weakness or strength. For example, in first grade it was the extremely low performance on Following Directions that helps explain why class 1 performed below the 25th percentile on reading comprehension when this class had similar performance to two other classes on Vocabulary Pairs and Word Reading. Intervention for the 11 students in Class 1 would need to take into

account their very low skill in listening to and remembering concepts and directions by repeating directions and providing multiple practice opportunities to learn new concepts.

The latent classes of students in kindergarten through second grade with strong alphabetic skills generally had higher reading outcomes. Academic language strengths did not appear to compensate for weak alphabetic skills in these early grades. However, reading intervention should not simply focus on alphabetic skills, because academic language skills account for a large proportion of variance in comprehension even in the primary grades (Foorman et al., 2015a; Catts et al., 1999; Kendeou, van den Broek, White, & Lynch, 2009; Muter, Hulme, Snowling, & Stevenson, 2004).

In the current investigation, the importance of language skills to reading comprehension became even more obvious by the upper elementary grades. In third and fifth grades, very low Vocabulary Knowledge scores were associated with performance on latent reading comprehension that was more than 1 SD below the mean. This latent class of students was similar to the “word callers” that Buly and Valencia (2002) found in their cluster analysis of fourth grade students who had failed the state reading test. However, in the fourth grade sample in this study, the close to average performance on the other language task—Syntactic Knowledge—helped to offset the effect of low Vocabulary Knowledge on reading comprehension. Knowledge of sentence use (i.e., syntax) and knowledge of the structure and meaning of words (i.e., vocabulary) are highly related and strongly predict reading comprehension in fourth through tenth grades (Foorman et al., 2015b). It is not surprising, then, that tasks that measure the understanding of discourse connectors in text, such as the FRA Syntactic Knowledge task, and the understanding of word meanings and structure, such as the FRA Vocabulary Knowledge task, should inform relations to reading comprehension more than Word Recognition did. This does not mean that intervention for struggling readers should ignore weaknesses in word identification. It simply means that interventions also need to build knowledge of the structure and meanings of words and of the linguistic devices for making text cohesive (e.g., Kieffer & Lesaux, 2012; Lawrence, Crosson, Paré-Blagoev, & Snow, 2015; Lesaux, Kieffer, Faller, & Kelley, 2010). Furthermore, the identification of this substantive group of low-performers in middle and high school grades highlights the need to provide differentiated instruction in the form of more intensive support for academic language skills that is often not provided in typical middle and high school classrooms.

The heterogeneity of profiles in the elementary grades compared to the uniformity of profiles in the secondary grades suggests that the elementary grades are an important time to intervene and prevent future widening of the gap in reading comprehension. In addition, it is crucial that differentiation be based on learning profiles derived from valid and reliable measures. Unreliable descriptions of students’ strengths and weaknesses can lead to inappropriate instruction even when based on authentic tasks such as oral reading (Denton et al., 2006). Even with reliable measures, metrics (e.g., oral reading) may not be on a vertical scale, making measurement of growth across grade levels or across forms invalid (Francis et al., 2008). Additionally, profiles may be invalid if they omit measures critical to defining the underlying construct, such as omitting language measures from the creation of reading profiles. Finally, unless learning profiles are linked to recognized outcomes, as was done in

this investigation and in Brasseur-Hock et al. (2011) and in Buly and Valencia (2002), the goals for differentiated instruction to achieve important reading outcomes may be less clear.

An important next step for the field is to test the results of the heterogeneous profiles from this exploratory latent profile analysis with confirmatory latent class analysis. Although the samples at each grade were not sufficiently large to conduct confirmatory latent class analyses, this investigation nonetheless serves as an important first step in verifying the existence of various groupings of students (e.g., the poor comprehenders of Catts et al., 2012) and validating the instructional utility of diagnostic profiles, which need to be conducted through future intervention studies. Further, research using factor mixture modeling could provide a more in-depth test of theoretical construct subgroupings.

Limitations

A major limitation of the current study is that it is cross-sectional rather than longitudinal. By conducting the LPA by grade and associating profiles with reading outcomes at each grade it is impossible to say whether a profile in kindergarten predicts later reading comprehension performance as the Catts et al. (2012) does. However, the replication of more heterogeneous profiles in the elementary grades compared to the secondary grades suggests that differentiating instruction based on the learning profiles should occur in the earlier grades. An additional limitation is that the profiles and their relations to reading outcomes were limited to the measures used. The measures used in this study are part of a universal screening battery. Since the purpose of universal screening is to provide a reliable score in a short amount of time, the construct validity of each individual task is not as high as gold standard clinical assessments of the same constructs and the construct coverage is not as extensive. Validity coefficients for the FRA are similar to other commercial screening assessments. Furthermore, this limitation is offset by use of a latent variable of reading comprehension and its strong relation to the single measures in the FRA reading and language measures. There is still quite a bit of debate in the research literature about the relative importance of different components of language and when fluency predicts unique variance above and beyond word recognition and language skills. The profile measures used in this study were chosen based on recently converging evidence regarding the important components of reading (Foorman et al., 2015a, 2015b; Language & Reading Research Consortium (LARCC), 2015). However, it is possible that other constructs (e.g., reading fluency) may produce another profile represented by a small number of students.

Conclusions

The idea of differentiating instruction based on students' learning profiles has gained popularity with the increased emphasis on data-based decision making (Fuchs & Fuchs, 2006; Tomlinson, 2007). Reading educators have been using informal reading inventories and surveys for decades to describe students' strengths and weaknesses. These descriptive approaches are useful at the classroom-level for differentiating instruction. However, when used for purposes of classification, placement, and monitoring growth, their psychometric limitations become apparent (e.g., Denton et al., 2006). Quantitative approaches to classifying and clustering students by skill profiles have become reliable and stable with the use of multiple measures to form latent classes and the introduction of statistical tests and

goodness-of-fit indices of Latent Class and Latent Profile Analysis to test hypotheses about the number of classes that exist in a population of students.

The Latent Profile Analysis conducted in this study with 7,752 students in kindergarten through tenth grade utilized a computer-adaptive battery of reading and language measures with strong psychometric properties. The resulting profiles related moderately to strongly with reading outcomes. Profiles in the secondary grades fell into a pattern of high, medium, and low. Profiles in the elementary grades were more heterogeneous, with the role of academic language becoming increasing apparent across the grades. The resulting profiles lend strong support to differentiating instruction by addressing students' academic language needs in addition to their word identification needs.

Acknowledgments

Funding

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through a subaward to Florida State University from Grant R305F100005 to the Educational Testing Service as part of the Reading for Understanding Initiative, and the National Institute of Child Health and Human Development from Grant P50HD052120. The opinions expressed are those of the authors and do not represent views of the Institute, the U.S. Department of Education, the Educational Testing Service, the National Institute of Child Health and Human Development or Florida State University.

References

- Argyle S (1989). Miscue analysis for classroom use. *Reading Horizons*, 29(2), 93–102.
- Arndt E, & Foorman B (2010). Second graders as spellers: What types of errors are they making? *Assessment for Effective Intervention*, 36(1), 57–67. 10.1177/1534508410380135
- Badian N, Duffy F, Als H, & McAnulty G (1991). Linguistic profiles of dyslexic and good readers. *Annals of Dyslexia*, 41, 221–245. [PubMed: 24233767]
- Benjamini Y, & Hochberg Y (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B Methodological*, 57(1), 289–300.
- Biancarosa C, & Snow CE (2006). *Reading next—A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York* (2nd ed.). Washington, DC: Alliance for Excellent Education.
- Brasseur-Hock I, Hock M, Kieffer M, Biancarosa G, & Deshler D (2011). Adolescent struggling readers in urban schools: Results of a latent class analysis. *Learning and Individual Differences*, 21, 438–452. 10.1016/j.lindif.2011.01.008
- Brimijoin K, Marquisee E, & Tomlinson C (2003). Using data to differentiate instruction. *Educational Leadership*, 60(5), 70–72.
- Buly M, & Valencia S (2002). Below the bar: Profiles of students who fail state reading assessments. *Education Evaluation & Policy Analysis*, 24(3), 219–239. 10.3102/01623737024003219
- Cain K, Patson N, & Andrews L (2005). Age- and ability-related differences in young readers' use of conjunctions. *Journal of Child Language*, 32, 877–892. 10.1017/s030500090500704 [PubMed: 16429715]
- Carrow-Woolfolk E (2008). *Comprehensive Assessment of Spoken Language*. (n.p.): Western Psychological Services.
- Catts H, Adlof S, & Weismer S (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, 49, 278–293. 10.1044/1092-4388(2006/023)
- Catts H, Compton D, Tomblin JB, Bridges M (2012). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology*, 104(1), 166–181. 10.1037/a0025323

- Catts H, Fey M, Zhang X, & Tomblin JB (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading*, 3(4), 331–361. 10.1207/s1532799xssr0304.2
- Conner CM, Morrison FJ, Fishman BJ, Schatschneider C & Underwood P (2007). The early years: Algorithm-guided individualized reading instruction. *Science*, 315 (5811) 464–465. doi:10.1126/science.1134513 [PubMed: 17255498]
- Denton C, Ciancio D, Fletcher J (2006). Validity, reliability, and utility of the Observation Survey of Early Literacy Achievement. *Reading Research Quarterly*, 41(1), 8–34. 10.1598/RRQ.41.1.1
- Dunn L, & Dunn D (2007). *Peabody Picture Vocabulary Test-4*. San Antonio, TX: Pearson.
- Florida Department of Education. (2013). *Florida Statewide Assessments 2013 Yearbook*. (n.p.: Author).
- Foorman B, Herrera S, Petscher Y, Mitchell A, & Truckenmiller A (2015). The Structure of Oral Language and Reading and Their Relation to Comprehension in grades kindergarten through grade 2. *Reading and Writing*, 28(5), 655–681. 10.1007/s11145-015-9544-5 [PubMed: 27660395]
- Foorman B, Koon S, Petscher Y, Mitchell A, & Truckenmiller A (2015). Examining General and Specific Factors in the Dimensionality of Oral Language and Reading in 4th-10th Grades. *Journal of Educational Psychology*, 107(3), 884–899. 10.1037/edu0000026 [PubMed: 26346839]
- Foorman B, Petscher Y, & Bishop MD (2012). The Incremental Variance of Morphological Knowledge to Reading Comprehension in Grades 3–10 Beyond Prior Reading Comprehension, Spelling, and Text Reading Efficiency. *Learning and Individual Differences*. 22(6), 792–798. 10.1016/j.lindif.2012.07.009
- Foorman B, Petscher Y, & Schatschneider C (2015c). Florida Center for Reading Research (FCRR) Reading Assessments (FRA) Kindergarten to Grade 2 Technical Manual. Retrieved from <http://www.fcrr.org/for-researchers/fra.asp>
- Foorman B, Petscher Y, & Schatschneider C (2015d). Florida Center for Reading Research (FCRR) Reading Assessments (FRA) Grade 3 Through 12 Technical Manual. Retrieved from <http://www.fcrr.org/for-researchers/fra.asp>
- Foorman B, & Wanzek J (2015). Classroom Reading Instruction for all students In Jimerson SR, Burns MK, & VanDerHeyden AM (Eds.), *The Handbook of Response to Intervention: The Science and Practice of Multi-Tiered Systems of Support* (2nd ed.). New York, NY: Springer Science.
- Francis D, Fletcher J, Stuebing K, Lyon GR, Shaywitz B, & Shaywitz S (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities*, 38, 98–108. 10.1177/00222194050380020101 [PubMed: 15813593]
- Francis D, Santi K, Barr C, Fletcher J, Varisco A, & Foorman B (2008). Form effects on the estimation of students' oral reading fluency using DIBELS. *Journal of School Psychology*, 46, 315–342. 10.1016/j.jsp.2007.06.003 [PubMed: 19083362]
- Fuchs D, & Fuchs L (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, 41(1), 93–99. 10.1598/RRQ.41.1.4
- Garcia JR, & Cain K (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, 84(1), 74–111. 10.3102/0034654313499616
- Geva E, & Ryan E (1985). Use of conjunctions in expository texts by skilled and less skilled readers. *Journal of Literacy Research*, 17(4), 331–346. 10.1080/10862968509547549
- Graesser A, McNamara D, Louwerse M, & Cai Z (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruction, & Computers*, 36, 193–202. 10.3758/BF03195564
- Brace Harcourt (2003). *Stanford Achievement Test* (10th Ed.). San Antonio, TX: Author.
- Assessment Harcourt (2004). *Stanford Technical data report*. San Antonio, TX: Author.
- Halpin P, & Kieffer M (2015). Describing profiles of instructional practice: A new approach to analyzing classroom observation data. *Educational Researcher*, 44(5), 263–277. 10.3102/0013189X15590804
- Heritage M (2010). *Formative assessment: Making it happen in the classroom*. Thousand Oaks, CA: Corwin Press.

- Holliman A, Hurry J, & Bodman S (2014). Children's reading profiles on exiting Reading Recovery programme: do they predict sustained progress? *Journal of Research in Reading*, 39(1), 1–18. 10.1111/1467-9817.12041
- Justice L, Logan J, Kaderavek J, Schmitt MB, Tompkins V, & Barlett C (2015). Empirically based profiles of the early literacy skills of children with language impairment in early childhood special education. *Journal of Learning Disabilities*, 48(5), 482–494. 10.1177/0022219413510179 [PubMed: 24232733]
- Kendeou P, van den Brock P, White M, & Lunch J (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology*, 101, 765–778. 10.1037/a0015956
- Kieffer M, & Lesaux N (2012). Effects of academic language instruction on relational and syntactic aspects of morphological awareness for sixth graders from linguistically diverse backgrounds. *Elementary School Journal*, 112(3), 519–545. 10.1086/663299
- Language and Reading Research Consortium (LARRC) (2015). The dimensionality of language in young children. *Child Development*, 86, 1948–1965. 10.1111/cdev.12450 [PubMed: 26509742]
- Language and Reading Research Consortium (LARCC) (in press). Oral language and listening comprehension: Same or different constructs? *Journal of Speech, Language, and Hearing Research*.
- Lawrence J, Crosson A, Paré-Blagoev, & Snow C (2015). Word generation randomized trial: Discussion mediates the impact of program treatment on academic word learning. *American Educational Research Journal*, 52(4), 750–786. 10.3102/0002831215579485
- Lesaux N, Kieffer M, Faller S, & Kelley J (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for sixth graders in urban middle schools. *Reading Research Quarterly*, 45(2), 196–228. 10.1598/RRQ.45.2.3
- Logan J, & Petscher Y (2010). School profiles of at-risk student concentration: differential growth in oral reading fluency. *Journal of School Psychology*, 48(2), 163–186. 10.1016/j.jsp.2009.12.002 [PubMed: 20159224]
- Lord F, & Novick M (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Lubke G & Muthén B (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, 10, 21–39. 10.1037/1082-989x.10.1.21 [PubMed: 15810867]
- McKenna M, & Stahl K (2009). *Assessment for Reading Instruction* (2nd ed.). New York, NY: Guilford Press.
- Morris R, Steubing K, Fletcher J, Shaywitz S, Lyon GR, Shankweiler D, Katz L, & Francis D (1998). Subtypes of reading disability: Variability around a phonological core. *Journal of Educational Psychology*, 90(3), 347–373.
- Muter V, Hulme C, Snowling M, & Stevenson J (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology*, 40(5), 665–681. 10.1037/0012-1649.40.5.665 [PubMed: 15355157]
- National Institute of Child Health and Human Development (2000). *National Reading Panel—Teaching children to read: Reports of the subgroups* (NIH Pub. No. 00–4754). Washington, D.C.: U.S. Department of Health and Human Services.
- Samuelsen K, & Raczynski K (2013). Latent class/profile analysis In Petscher Y, Schatschneider C, & Compton D (Eds.), *Applied quantitative analysis in education and the social sciences* (pp. 304–328). New York, NY: Routledge.
- Schleppegrell M (2012). Academic language in teaching and learning: Introduction to special issue. *Elementary School Journal*, 112, 409–418.
- Semel E, Wigg E, & Secord W (2003). *The Clinical Evaluation of Language Fundamentals*, 4th edition: Examiner's Manual. San Antonio, TX: Pearson.
- Shankweiler D, Crain S, Katz L, Fowler A, Liberman A, Brady S, Thornton R, Lundquest E, Dreyer L, Fletcher J, Steubing K, Shaywitz S, & Shaywitz B (1995). Cognitive profiles of reading disabled children: Comparison of language skills in phonology, morphology, and syntax. *Psychological Science*, 48(2), 163–186.

- Stanovich KE (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 360–407. doi:10.1598/rrq.21.4.1
- Stanovich K, & Siegel L (1994). Phenotypic performance profile of children with reading disabilities: Regression-based test of phonological-core variable-difference model. *Journal of Educational Psychology*, 86(1), 24–53. 10.1037/0022-0663.86.1.24
- Tomlinson C (1999). Mapping a route toward differentiated instruction. *Educational Leadership*, 57(1), 12–16.
- Tomlinson C (2007). Learning to love assessment. *Educational Leadership*, 65(4), 8–13.
- Torgesen JK, Wagner RK, & Rashotte CA (2012). *Test of Word Reading Efficiency-2*. Austin, TX: Pro-Ed.
- Uccelli P, Galloway E, Barr C, Meneses A, & Dobbs C (2015). Beyond vocabulary: Exploring cross-disciplinary academic-language proficiency and its association with reading comprehension. *Reading Research Quarterly*, 50(3), 337–356. 10.1002/rrq.104
- Van Silfhout G, Evers-Vermeul J, Mak W, & Sanders T (2014). Connectives and layout as processing signals: How textual features affect students' processing and text representation. *Journal of Educational Psychology*. 106(4), 1036–1048. 10.1037/a0036293
- Vellutino F, Scanlon D, Sipay E, Small S, Pratt A, Chen R, & Denkla M (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, 88, 601–638.
- Vellutino F, Scanlon D, Small S, & Fanuele D (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities*, 39(2), 157–169. 10.1177/00222194060390020401 [PubMed: 16583795]
- Wagner RK, Herrera S, Spencer M, Quinn J (2014). Reconsidering the simple view of reading in an intriguing case of equivalent models: Commentary on Tunmer and Chapman (2012). *Journal of Learning Disabilities*, 48, 115–119. doi:10.1177/0022219414544544 [PubMed: 25095787]
- Woodcock R, McGrew K, & Mather N (2001). *Woodcock-Johnson-III Tests of Achievement*. Itasca, IL: Riverside.
- Zeno SM, Ivens SH, Millard RT, Duvvuri R (1995). *The educator's word frequency guide* NY: Touchstone Applied Science Associates, Inc.

Figure 1.

Clockwise from top left, profile plots reflect data for kindergarten, first, and second grades, respectively, for FRA measures of Vocabulary Pairs (VOC), Following Directions (FD), Phonological Awareness (PA), Letter Sounds (LS), Sentence Comprehension (SC), Word Reading (WR), and Spelling (Spell). The lines represent distinct emergent profiles (i.e., c1, c2, c3, etc.).

Figure 2.

Clockwise from top left, profile plot reflect data for third, fourth, fifth, and eighth, respectively, for FRA measures of Vocabulary Knowledge Task (VKT), Word Recognition Task (WRT), and Syntactic Knowledge Task (SKT). The lines represent distinct emergent profiles (i.e., c1, c2, c3, etc.).

Table 1.

Number of participants by grade and percentages for gender, race/ethnicity, English learner (EL), and free-and-reduced-lunch (FRL) status.

Grade	N	Gender		Race/Ethnicity			Status	
		Female	Black	Hispanic	White	Other	EL	FRL
K	422	50.50	22.98	30.94	39.80	6.28	18.11	65.34
1	989	51.60	22.69	30.68	40.36	6.26	18.40	65.68
2	884	50.20	22.56	30.43	40.53	6.48	16.18	64.81
3	607	52.16	22.91	30.36	40.29	3.86	7.36	64.97
4	584	51.36	21.95	29.91	41.49	3.93	6.91	63.18
5	659	51.65	22.21	29.52	41.77	3.84	8.54	62.95
6	770	48.84	22.75	29.22	41.73	3.71	9.63	62.39
7	696	48.55	22.73	29.05	42.10	3.56	8.87	60.53
8	629	50.04	22.32	29.00	42.86	3.37	8.72	58.63
9	704	50.63	22.60	28.27	43.20	3.42	5.40	54.13
10	808	50.61	22.73	27.95	43.46	3.21	6.21	50.72
Overall	7752	50.31	22.55	29.10	42.18	3.59	10.39	59.69
State	2.2 million	50.00	23.00	30.00	41.00	6.00	8.00	60.00

Note. The official enrollment for Florida students in kindergarten through grade 10 during the 2012/13 school year was 2,259,791.

Table 2.

Descriptive statistics and correlations for measures for kindergarten, grade 1, and grade 2

Measure	1	2	3	4	5	6
<u>Kindergarten</u>						
1. SESAT	1.00	.38	.46	.58	.51	.28
2. VOC	--	1.00	.45	.32	.28	.41
3. FD	--	--	1.00	.42	.33	.54
4. PA	--	--	--	1.00	.48	.33
5. LS	--	--	--	--	1.00	.27
6. SC	--	--	--	--	--	1.00
<i>n</i>	321	422	422	422	422	422
Raw <i>M</i>	448.02	499.95	499.99	500.14	500.33	499.99
Raw <i>SD</i>	46.60	97.60	99.88	98.45	99.20	99.52
Range (L, H)	(349.00, 565.00)	(178.00, 785.00)	(218.00, 720.00)	(238.00, 727.00)	(234.00, 665.00)	(119.18, 702.45)
<u>Grade 1</u>						
1. SAT-10	1.00	.58	.51	.75		
2. VOC	--	1.00	.51	.55		
3. FD	--	--	1.00	.41		
4. WR	--	--	--	1.00		
<i>n</i>	989	892	979	237		
Raw <i>M</i>	589.53	500.00	500.00	500.00		
Raw <i>SD</i>	49.37	100.00	100.00	100.00		
Range (L, H)	(443.00, 666.00)	(125.00, 714.00)	(79.00, 703.00)	(177.00, 676.00)		
<u>Grade 2</u>						
1. SAT-10	1.00	.55	.49	.62	.58	
2. VOC	--	1.00	.39	.44	.42	
3. FD	--	--	1.00	.35	.36	
4. SPELL	--	--	--	1.00	.77	
5. WR	--	--	--	--	1.00	
<i>n</i>	884	846	871	852	235	
Raw <i>M</i>	618.87	500.00	500.00	500.00	500.00	
Raw <i>SD</i>	43.30	100.00	100.00	100.00	100.00	
Range (L, H)	(489.00, 726.00)	(90.00, 695.00)	(147.00, 744.00)	(189.00, 798.00)	(230.00, 666.00)	

Note. All correlations are significant at the .01 level. SESAT= Stanford Early Scholastic Achievement Test, Form A; VOC=Vocabulary Pairs; FD=Following Directions; PA=Phonological Awareness; LS=Letter Sounds; SC=Sentence Comprehension; SAT-10=Stanford Achievement Test, 10th edition (Reading Comprehension); WR=Word Reading; SPELL=Spelling.

Table 3.

Descriptive statistics and correlations for measures for grades 3–5 and grade 8

Measure	1	2	3	4	5	6
<u>Grade 3 (n = 607)</u>						
1. VKT	1.00	.33	.31	.57	.53	.54
2. WRT	--	1.00	.29	.41	.39	.35
3. SKT	--	--	1.00	.40	.38	.38
4. RCT	--	--	--	1.00	.76	.77
5. SAT-10	--	--	--	--	1.00	.81
6. FCAT	--	--	--	--	--	1.00
Raw <i>M</i>	498.97	499.24	501.61	381.98	644.14	201.00
Raw <i>SD</i>	98.24	96.56	97.23	66.40	44.47	21.80
Range (L, H)	(212.00, 810.00)	(223.00, 956.00)	(232.00, 812.00)	(260.00, 571.00)	(522.00, 740.00)	(140.00, 260.00)
<u>Grade 4 (n = 587)</u>						
1. VKT	1.00	.29	.36	.43	.43	.42
2. WRT	--	1.00	.35	.45	.36	.39
3. SKT	--	--	1.00	.53	.50	.54
4. RCT	--	--	--	1.00	.72	.78
5. SAT-10	--	--	--	--	1.00	.75
6. FCAT	--	--	--	--	--	1.00
Raw <i>M</i>	500.16	500.44	500.17	465.15	655.41	214.14
Raw <i>SD</i>	93.94	94.31	95.51	64.21	40.20	21.70
Range (L, H)	(86.00, 877.00)	(246.00, 813.00)	(185.00, 954.00)	(339.00, 654.00)	(522.00, 761.00)	(154.00, 269.00)
<u>Grade 5 (n = 659)</u>						
1. VKT	1.00	.38	.46	.59	.55	.58
2. WRT	--	1.00	.35	.42	.39	.45
3. SKT	--	--	1.00	.59	.56	.62
4. RCT	--	--	--	1.00	.75	.81
5. SAT	--	--	--	--	1.00	.81
6. FCAT	--	--	--	--	--	1.00
Raw <i>M</i>	500.11	499.72	499.64	478.22	665.80	219.86
Raw <i>SD</i>	95.25	94.74	96.03	90.06	37.16	21.88
Range (L, H)	(94.00, 759.00)	(222.00, 888.00)	(192.00, 810.00)	(267.00, 680.00)	(554.00, 777.00)	(161.00, 277.00)
<u>Grade 8 (n = 629)</u>						
1. VKT	1.00	.33	.38	.52	.46	.56
2. WRT	--	1.00	.51	.42	.38	.44
3. SKT	--	--	1.00	.56	.51	.57
4. RCT	--	--	--	1.00	.67	.75
5. SAT	--	--	--	--	1.00	.71
6. FCAT	--	--	--	--	--	1.00
Raw <i>M</i>	499.99	500.21	500.24	558.06	685.13	231.26
Raw <i>SD</i>	100.00	79.59	85.32	138.82	32.81	21.72

Measure	1	2	3	4	5	6
Range (L, H)	(117.00, 921.00)	(198.00, 797.00)	(126.00, 830.00)	(337.00, 927.00)	(593.00, 801.00)	(175.00, 296.00)

Note. All correlations are significant at the .01 level. VKT= Vocabulary Knowledge Task; WRT=Word Recognition Task; SKT=Syntactic Knowledge Task; RCT=Reading Comprehension Task; SAT-10=Stanford Achievement Test, 10th edition (Reading Comprehension); FCAT=Florida Comprehensive Assessment Test (Reading Comprehension).

Table 4.

Latent profile model fit for kindergarten through grade 5 and grade 8

Grade	Profiles	Parameters	LL	AIC	aBIC	–2LL
K	2	19	–3255.01	6548.01	6624.87	
	3	22	–2681.62	5407.25	5496.24	1146.77*
	4	28	–2653.41	5362.82	5476.08	56.42*
	5	34	–2629.75	5357.51	5465.04	47.32*
	6	40	–2618.44	5316.88	5458.68	22.63*
1	2	10	–2818.26	5656.52	5705.49	
	3	14	–2785.14	5598.28	5666.84	66.24*
	4	18	–2768.46	5572.92	5661.01	33.36*
	5	22	–2752.99	5549.98	5657.71	30.94*
	6	26	–2743.43	5546.86	5674.17	19.12*
2	2	13	–3768.29	7562.59	7624.79	
	3	18	–3697.13	7430.26	7516.38	142.33*
	4	23	–3669.02	7384.03	7494.08	56.22*
	5	28	–3655.54	7367.07	7501.04	26.96*
	6	33	–3642.95	7355.89	7513.78	25.17*
3	2	10	–2202.90	4425.81	4438.14	
	3	14	–2173.78	4375.56	4392.83	58.24*
	4	18	–2154.42	4344.83	4367.04	38.73*
	5	22	–2129.87	4303.74	4330.88	49.10*
	6	26	–2104.72	4261.43	4293.51	50.30*
4	2	10	–2166.75	4353.49	4365.49	
	3	14	–2140.35	4308.69	4325.50	52.80*
	4	18	–2112.79	4261.58	4283.18	55.12*
	5	22	–2097.77	4239.54	4265.95	30.04*
	6	26	–2087.22	4226.44	4257.65	21.10*
5	2	10	–2451.39	4922.79	4935.95	
	3	14	–2405.92	4839.83	4858.25	90.96*
	4	18	–2383.61	4803.22	4826.91	44.61*
	5	22	–2363.13	4770.25	4799.20	40.97*
	6	26	–2353.85	4759.70	4793.91	18.55*
8	2	10	–1996.36	4012.71	4025.41	
	3	14	–1954.42	3936.83	3954.60	83.88*
	4	18	–1924.65	3885.30	3908.15	59.53*
	5	22	–1902.37	3848.73	3876.65	44.57*
	6	26	–1885.47	3822.95	3855.95	33.78*

Note. LL = log likelihood, AIC = Akaike Information Criteria, aBIC = sample adjusted Bayes Information Criteria, –2LL = log likelihood ratio test. Values in bold represent final selected class.

* $p < .001$.

Table 5.

General linear model contrasts among latent profile classes for Kindergarten through grade 2

Grade	Comparison	Estimate	S.E.	t-value	<i>p</i>	Critical <i>p</i>	Hedge's <i>g</i>
K	c1 vs c2	-46.35	15.14	-3.06	.002	.027	-1.56
	c1 vs c3	-53.60	8.13	-6.59	<.001	.003	-1.33
	c1 vs c4	-23.38	8.43	-2.77	.006	.030	-0.90
	c1 vs c5	-29.11	11.52	-2.53	.012*	.006	-1.05
	c1 vs c6	-100.16	9.13	-10.97	<.001	.007	-2.26
	c2 vs c3	-7.25	13.57	-0.53	.593	–	-0.18
	c2 vs c4	22.96	13.75	1.67	.096	–	0.93
	c2 vs c5	17.24	15.83	1.09	.277	–	0.68
	c2 vs c6	-53.82	14.19	-3.79	<.001	.010	-1.13
	c3 vs c4	30.21	5.12	5.90	<.001	.013	0.84
	c3 vs c5	24.49	9.37	2.61	.009	.033	0.61
	c3 vs c6	-46.56	6.19	-7.52	<.001	.017	-1.06
	c4 vs c5	-5.73	9.63	-0.59	.553	–	-0.23
	c4 vs c6	-76.78	6.59	-11.66	<.001	.020	-2.15
	c5 vs c6	-71.05	10.24	-6.94	<.001	.023	-1.59
1	c1 vs c2	-116.87	12.29	-9.51	<.001	.005	-3.40
	c1 vs c3	-44.96	14.21	-3.16	.002	.045	-1.09
	c1 vs c4	-40.91	12.48	-3.28	<.001	.010	-1.17
	c1 vs c5	-81.80	12.26	-6.67	<.001	.015	-2.18
	c2 vs c3	71.92	7.76	9.27	<.001	.020	1.64
	c2 vs c4	75.97	3.72	20.41	<.001	.025	2.02
	c2 vs c5	35.17	2.90	12.12	<.001	.030	0.87
	c3 vs c4	4.06	8.05	0.50	.614	–	0.09
	c3 vs c5	-36.74	7.70	-4.77	<.001	.035	-0.80
	c4 vs c5	-40.80	3.60	-11.33	<.001	.040	-1.01
2	c1 vs c2	-64.92	5.92	-10.98	<.001	.003	-2.17
	c1 vs c3	-33.99	5.62	-6.05	<.001	.007	-1.32
	c1 vs c4	-72.39	5.24	-13.82	<.001	.010	-2.51
	c1 vs c5	-104.59	5.24	-19.95	<.001	.013	-3.55
	c1 vs c6	-43.96	6.69	-6.57	<.001	.017	-1.72
	c2 vs c3	30.93	4.35	7.11	<.001	.020	0.99
	c2 vs c4	-7.47	3.84	-1.95	<.001	.023	-0.22
	c2 vs c5	-39.67	3.84	-10.32	<.001	.027	-1.16
	c2 vs c6	20.96	5.66	3.70	<.001	.030	0.67
	c3 vs c4	-38.40	3.40	-11.40	<.001	.033	-1.27
	c3 vs c5	-70.60	3.38	-20.91	<.001	.037	-2.29
	c3 vs c6	-9.96	5.36	-1.86	<.001	.040	-0.37
	c4 vs c5	-32.21	2.60	-11.98	<.001	.043	-0.96

Grade	Comparison	Estimate	S.E.	t-value	<i>p</i>	Critical <i>p</i>	Hedge's <i>g</i>
	c4 vs c6	28.43	4.95	5.74	<.001	.047	0.94
	c5 vs c6	60.64	4.96	12.24	<.001	.050	1.98

* *p* value was not statistically significant after applying linear step-up correction (i.e., $p > \text{Critical } p$). All remaining *p* values < .05 were statistically significant after applying linear step-up correction (i.e., $p < \text{Critical } p$).

Table 6.

Latent profiles of reading and language variables for Kindergarten through grade 2 students

Grade	Class	Description	% (N)
K	1	Low on all variables	7% (32)
	2	Average on VOC, FD, LS; above average on PA; very low on SC	2% (8)
	3	Above average on all variables except VOC	42% (117)
	4	Average on language variables; below average on PA and LS	23% (97)
	5	Average on PA and LS; below average on language variables	7% (28)
	6	High on all variables	19% (80)
1	1	Very low on FD; below average on VOC and WR	1% (11)
	2	High on all variables	35% (346)
	3	High on FD; below average on VOC and WR	3% (29)
	4	Low on all variables	17% (175)
	5	Average on all variables	43% (428)
2	1	Very low on all variables	5% (43)
	2	Above average on WR and Spell; below average on VOC and FD	10% (92)
	3	Low on all variables	15% (132)
	4	Above average on VOC and FD; below average on WR and Spell	32% (286)
	5	High on all variables, especially Spell and WR	32% (282)
	6	Above average on VOC and FD; very low on WR and Spell	5% (49)

Note. % (N) = percentage and number of students. The abbreviations for the K-2 FRA are: VOC = Vocabulary Pairs; FD = Following Directions; PA = Phonological Awareness; LS = Letter Sounds; SC = Sentence Comprehension; WR = Word Reading; Spell = Spelling.

Table 7.

General linear model contrasts among latent profile classes in grades 3–5 and grade 8

Grade	Comparison	Estimate	S.E.	t-value	<i>p</i>	Critical <i>p</i>	Hedge's <i>g</i>
3	c1 vs c2	−0.61	0.27	−2.31	.021	.040	−0.53
	c1 vs c3	−0.99	0.19	−5.30	<.001	.005	−1.04
	c1 vs c4	−2.43	0.23	−10.71	<.001	.010	−2.94
	c1 vs c5	−2.05	0.25	−8.23	<.001	.015	−2.47
	c2 vs c3	−0.37	0.20	−1.86	.063	–	−0.52
	c2 vs c4	−1.81	0.24	−7.64	<.001	.020	−2.07
	c2 vs c5	−1.44	0.26	−5.55	<.001	.025	−1.67
	c3 vs c4	−1.44	0.14	−10.16	<.001	.030	−1.45
	c3 vs c5	−1.07	0.18	−6.07	<.001	.035	−1.21
	c4 vs c5	0.38	0.22	1.73	.085	–	0.25
4	c1 vs c2	−1.45	0.13	−10.77	<.001	.005	−1.40
	c1 vs c3	0.12	0.32	0.38	.703	–	0.25
	c1 vs c4	−1.91	0.17	−11.07	<.001	.010	−2.62
	c1 vs c5	−0.37	0.16	−2.28	.023	.040	−0.54
	c2 vs c3	1.57	0.31	5.12	<.001	.015	1.56
	c2 vs c4	−0.46	0.14	−3.25	.001	.035	−0.68
	c2 vs c5	1.08	0.13	8.37	<.001	.020	0.92
	c3 vs c4	−2.03	0.33	−6.25	<.001	.025	−2.83
	c3 vs c5	−0.49	0.32	−1.54	.125	–	−0.75
	c4 vs c5	1.54	0.17	9.16	<.001	.030	1.80
5	c1 vs c2	−0.72	0.30	−2.44	.015	.050	−1.17
	c1 vs c3	−1.43	0.26	−5.60	<.001	.005	−2.01
	c1 vs c4	−2.55	0.26	−9.93	<.001	.010	−3.36
	c1 vs c5	−3.54	0.28	−12.51	<.001	.015	−5.65
	c2 vs c3	−0.71	0.16	−4.40	<.001	.020	−1.05
	c2 vs c4	−1.83	0.16	−11.15	<.001	.025	−2.40
	c2 vs c5	−2.82	0.20	−13.93	<.001	.030	−4.37
	c3 vs c4	−1.12	0.07	−16.00	<.001	.035	−1.31
	c3 vs c5	−2.11	0.14	−15.38	<.001	.040	−2.64
	c4 vs c5	−0.98	0.14	−7.00	<.001	.045	−1.34
8	c1 vs c2	−1.46	0.08	−18.16	<.001	.012	−1.61
	c1 vs c3	1.10	0.18	6.07	<.001	.033	1.42
	c2 vs c3	2.56	0.19	13.42	<.001	.050	3.53

Note. All *p* values < .05 were statistically significant after applying linear step-up correction (i.e., *p* < Critical *p*).

Table 8.

Latent profiles of reading and language variables for grade 3–5 and 8 students

Grade	Class	Description	% (N)
3	1	Low on all variables, very low on VKT	5% (28)
	2	Low on all variables, very low on WRT	4% (25)
	3	Average on all variables	78% (472)
	4	High on all variables; higher on WRT	8% (49)
	5	High on all variables; higher on VKT	5% (33)
4	1	Low on all variables	12% (70)
	2	Average on all variables	53% (309)
	3	Low on all variables; very low on VKT	1% (8)
	4	High on all variables; very high on VKT	9% (52)
	5	Average on VKT; low on WRT and SKT	25% (145)
5	1	Low on WRT and SKT; very low on VKT	1% (8)
	2	Very low on all variables	4% (26)
	3	Low on all variables	53% (350)
	4	Medium on all variables	35% (231)
	5	Very high on all variables	7% (44)
8	1	Low on all variables	72% (451)
	2	High on all variables	25% (158)
	3	Very low on all variables, especially VKT	3% (20)

Note. % (N) = percentage and number of students. The abbreviations for the 3–10 FRA are: VKT = Vocabulary Knowledge Task; WRT = Word Recognition Task; SKT = Syntactic Knowledge Task.