Early Reading Skill Profiles in Typically Developing and At-Risk First Grade Readers to Inform Targeted Early Reading Instruction

Ryan P. Grimm[a]
University of California, Davis

Emily J. Solari[a]
University of California, Davis

Nancy S. McIntyre[a]
University of California, Davis

Carolyn A. Denton[b]
Children's Learning Institute
University of Texas Health Science Center Houston

Corresponding Author:
[a]Ryan P. Grimm, Ph.D.
School of Education
University of California, Davis
1 Shields Ave.
Davis, CA 95616
United States
rpgrimm@ucdavis.edu

[a]Emily J. Solari, Ph.D.
School of Education
University of California, Davis
1 Shields Ave.
Davis, CA 95616
United States
530-312-7172 phone and fax
ejsolari@ucdavis.edu

[a]Nancy S. McIntyre
School of Education
University of California, Davis
1 Shields Ave.
Davis, CA 95616
United States
nsmcintyre@ucdavis.edu

[b]Carolyn A. Denton, Ph.D.
Children’s Learning Institute
University of Texas Health Science Center Houston
7000 Fannin, UCT 2443
Houston, TX 77030
United States
(713) 500-3891 Phone
(713) 500-0315 Fax
Carolyn.A.Denton@uth.tmc.edu

Author Note

This work was supported by the Institute of Education Sciences, U.S. Department of Education, grant R324A150091 to the Regents of the University of California. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Institute of Education Sciences or the U.S. Department of Education.

Published in Journal of School Psychology v69 p111-126 May 2018

Abstract

This study identified distinct, homogeneous latent profiles of at-risk (*n* = 141) and not at-risk (*n* = 149) first grade readers. Separate latent profile analyses were conducted with each subgroup using measures of phonological awareness, decoding, linguistic comprehension, and oral reading fluency. This study also examined which measures best differentiated the latent profiles. Finally, we examined differences on two measures of reading comprehension as a function of profile membership. Results showed two latent profiles of at-risk students and three latent profiles of not at-risk students. Latent profiles were generally rank ordered with regard to achievement across measures. However, the higher performing at-risk profile and the lowest performing not at-risk profile were nearly identical across measures. Phonological awareness and decoding measures were best at differentiating latent profiles, but linguistic comprehension was also important for the lowest performing students. Oral reading fluency was limited to distinguishing the highest achieving students from the other profiles, and did not perform well with the lower achieving profiles. Most of the pairwise comparisons of reading comprehension scores were consistent across measures, but the nearly identical profiles showed a significant difference on only one reading comprehension measure. Implications for identifying at-risk first grade readers and designing targeted early reading interventions for at-risk students are discussed.

Keywords: reading disability, reading comprehension, reading intervention, latent profile analysis

Early Reading Skill Profiles in Typically Developing and At-Risk First Grade Readers to Inform Targeted Early Reading Instruction

In order to comprehend written text, an individual must weave together several reading subcomponent skills; early identification and remediation of these subcomponent skills is essential to avoid later reading comprehension difficulties. Difficulty acquiring the necessary sub-skills can lead to struggles with reading comprehension that can persist throughout a student's academic career (Cain & Oakhill, 2011; Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Torgesen et al., 2001). In a preventative model, early identification of struggling readers allows educational practitioners to provide intervention that may ameliorate later reading difficulties or disabilities. The purpose of this study was twofold. First, a two-step process was followed to identify and screen children who were at-risk for reading difficulties; we utilized a combination of teacher referral and researcher-led screening. Next, once an appropriate at-risk group was identified, we empirically derived distinct latent profiles based on multiple reading subcomponent skills to compare the number and nature of latent profiles across an early at-risk group of students and a group not at-risk for reading difficulties. This procedure was utilized to demonstrate the heterogeneous nature of reading subcomponent skill development within a group considered to be globally at risk of reading difficulties to help inform targeted early intervention efforts. A comparison group of not at-risk readers was utilized as a reference point for typical reading development, and also to demonstrate the range of reading profiles at first grade entry.

Multi-tiered systems of support (MTSS) models are comprehensive frameworks providing increasingly intensive academic support. In preventative MTSS reading models there are typically three tiers of instruction, each providing increasing support for students who are most in need (Fletcher & Vaughn, 2009; Fuchs & Fuchs, 2006; Fuchs, Mock, Morgan, & Young, 2003); intervention need is based on student data. In this framework, all students receive an evidence-based core reading curriculum from the general education teachers, this is also called Tier 1. When Tier 1 is deemed insufficient, based on student data, Tier 2 provides additional support, usually in the form of small-group interventions aimed at ensuring that students at-risk for reading difficulties meet grade-level benchmarks. Students with inadequate response to sufficiently intensive, evidence-based Tier 2 interventions receive more individualized and intensive Tier 3 interventions. While classroom teachers are often able to intuitively identify struggling students when their performance is starkly different from higher achieving peers (Begeny, Eckert, Montarello, & Storie, 2008; Gerber, 2005), the appropriate use of assessments enables a more nuanced picture of students' abilities and targeted intervention needs. This can inform the intervention design used within MTSS and may provide a more efficient intervention protocol. For instance, administering assessments in multiple reading domains may reveal targeted intervention needs. In terms of assessing early reading skills development, practitioners often focus on precursors of word reading such as phonological processing and letter and sound knowledge, as well as word level skills (decoding) and oral reading fluency. Linguistic comprehension has been shown to explain variance in reading comprehension in early elementary (Hoover & Gough, 1990; Kendeou, van den Broek, White, & Lynch, 2009; Storch & Whitehurst, 2002; Tunmer & Chapman, 2012), however it is not commonly assessed in early grades to determine risk status for reading. Given empirical evidence suggesting reading comprehension is a function of both decoding skills and linguistic comprehension (Hoover & Gough, 1990; Joshi & Aaron, 2000; Kendeou, van den Broek, et al., 2009; Ouellette & Beers, 2010; Tunmer & Chapman, 2012), it would seem reasonable to suggest that early readers would be screened for risk in both subcomponent skills.

The Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990) has been proposed as a useful way to categorize reading comprehension subcomponent skills, suggesting that word decoding and listening comprehension are the most significant predictors of reading comprehension. The Simple View of Reading has been utilized in two ways to help explain the development of reading comprehension. First, the framework has been used across multiple grade levels to explain the variance in reading comprehension performance (Adlof, Catts, & Lee, 2010; Adlof, Catts, & Little, 2006; Kendeou, Savage, & van den Broek, 2009; Kershaw & Schatschneider, 2010; Nation, Cocksey, Taylor, & Bishop, 2010; Tunmer & Chapman, 2012). It has also been used to describe subgroups of readers who struggle to develop the necessary sub-skills for successful reading comprehension (Catts, Adlof, & Weismer, 2006; Catts, Hogan, & Fey, 2003; Nation et al., 2010). It is imperative that school-based practitioners are able to identify students who struggle with early reading so that targeted reading intervention can be implemented.

Reading Sub-Skills

Several studies have used reading sub-skills (e.g., phonological processing, word reading, linguistic comprehension, and oral reading fluency) to differentiate readers (Catts, Compton, Tomblin, Bridges, 2012; Catts, Adlof, & Weismer, 2006; Catts, Fey, Zhang, & Tomblin, 1999; Nation et al., 2010). Other researchers have investigated how specific sub-skills predict later word reading and reading comprehension (Adlof et al., 2006; Hoover & Gough, 1990; Kendeou et al, 2009; Kershaw & Schatschneider, 2012; Roth, Speece, & Cooper, 2002; Silverman, Speece, Harring, & Richey, 2013; Storch & Whitehurst, 2002; Wagner & Torgesen, 1987). Since reading sub-skills can predict later reading comprehension, it follows these sub-skills can also be used to differentiate struggling readers from typically developing readers. Moreover, they may be used to identify distinct subgroups of struggling and typically developing readers. Many previous studies have treated sub-skills as independent contributors to reading comprehension, but few have examined how early reading profiles of sub-skills predict reading comprehension achievement. Examining profiles of reading sub-skills would enable researchers and practitioners to identify specific relative strengths and weaknesses that can inform tailored interventions.

Phonological Processing

Prior to word reading, children must develop phonological processing skills. Phonological processing is a multifaceted construct that involves using phonological information to process oral and written language (Catts, Fey, Zhang, & Tomblin, 1999; Wagner & Torgesen, 1987). The contribution of phonological processing to later word reading is well documented (e.g., Roth, Speece, & Cooper, 2002; Vellutino, Tunmer, Jaccard, & Chen, 2007; Wagner et al., 1997; Wagner & Torgesen, 1987). Since it is a precursor to word and text reading, phonological processing can be used to identify potentially struggling readers during the early elementary years. Though it is foundational to decoding, it is often mastered in early reading development, limiting its ability to distinguish readers as they progress.

Decoding and Reading Fluency

As students develop phonological processing skills, they begin to apply these to decoding words. When students decode words, they use their knowledge of letter-sound correspondence to read individual words. The significant relations between decoding and reading comprehension have been well-established in the literature (e.g., Chen & Vellutino, 1997; Georgiou, Das, & Hayward, 2009; Hoover & Gough, 1990; Joshi & Aaron, 2000). Accurate decoding is also a precursor to proficient oral reading fluency, which is an individual's ability to read text quickly and accurately. Proficient reading fluency reduces the cognitive load needed to read individual words and allows readers to shift attention to comprehending text (LaBerge & Samuels, 1974). Thus, it has been described as a "bridge" between decoding and reading comprehension (Pikulski & Chard, 2005) and is significantly related to reading comprehension (Chard et al., 2002; Fuchs, Fuchs, & Maxwell, 1988; Fuchs et al., 2001; Jenkins, Fuchs, van den Broek, Espin & Deno, 2003; Kim, Petscher, Schatschneider, & Foorman, 2010; Riedel, 2007; Silverman et al., 2013). As such, decoding is often targeted in early reading interventions (Foorman, Francis, Shaywitz, Shaywitz, & Fletcher, 1997; Foorman, Francis, Fletcher, Schatschneider, & Metha, 1998; Rashotte, MacPhee, & Torgesen, 2001) to improve reading fluency. Yet, interventions focused on decoding may not be sufficient to enable struggling readers to catch up to their peers. One reason may be that linguistic comprehension is also key to reading comprehension, even in the early elementary grades.

Linguistic Comprehension

Linguistic comprehension refers to a person's ability to understand spoken language. Extant research has found students who struggle with reading comprehension also struggle with linguistic comprehension (Catts et al., 2006; Catts et al., 1999; Nation, Adams, Bowyer-Crane, & Snowling, 1999; Nation et al., 2010). Thus, it makes sense to provide intervention that includes a linguistic comprehension component to struggling readers; emerging research has begun to demonstrate positive effects of linguistic comprehension intervention with respect to reading comprehension outcomes (Bowyer-Crane et al., 2008; Clarke, Snowling, Truelove, & Hulme, 2010). For example, Catts et al. (2006) found a profile of students who exhibited normal phonological processing skills, but experienced difficulties with linguistic and reading comprehension. The authors argued intervention for these students should target linguistic comprehension. Remediating phonological processing and decoding skills may lead to short-term gains in reading comprehension in early elementary, but, as texts become more complex, linguistic comprehension difficulties may impede later reading comprehension if they are not also addressed.

Differentiating Readers According to Reading Sub-Skills

Extant research has identified subgroups of readers using arbitrarily defined cutoff scores on one or more variables (or a composite variable) to differentiate readers based on reading sub skill strengths and weaknesses. Nation et al. (2010) conducted a study in which they examined the reading development of children between ages 5 and 8 who were accurate word readers, but poor comprehenders; poor comprehenders exhibited normal word reading development, but they experienced difficulties with linguistic comprehension. Both linguistic comprehension and reading comprehension difficulties were persistent across the ages. Since linguistic comprehension difficulties were apparent before the poor comprehenders acquired word reading skills, the authors concluded that early poor linguistic comprehension can place children at risk for later reading comprehension difficulties. Catts, Hogan, and Fey (2003) and Catts et al. (2006) went a step further by identifying multiple subgroups of poor readers. Catts et al. (2003) used measures of word reading and linguistic comprehension to subgroup poor readers in early elementary, defining poor readers as those who scored one standard deviation below the mean on a reading comprehension composite. They also separated the poor readers into four subgroups based on word reading and linguistic comprehension and found that linguistic comprehension skills in second grade were related to kindergarten measures of linguistic comprehension. Furthermore, they found the subgroups were moderately stable into fourth grade. However, the authors also reported that their sample did not form homogenous subgroups. They acknowledged this was related to their method of imposing a cutoff score and subgroup membership could shift depending on the cutoff score. As described below, the present study empirically derived subgroups rather than imposing a cutoff score.

Catts et al. (2006) created reading subgroups (i.e., poor comprehenders, poor decoders, and typical readers) from composite scores in reading comprehension and word recognition using the 25th and 40th percentiles. Specific deficits were defined as scoring below the 25th percentile in one domain, but above the 40th percentile in the other domain; students were examined at kindergarten, second, and fourth grade. Results showed the poor comprehenders scored significantly lower on a language comprehension composite compared to poor decoders and typical readers at all three timepoints. In contrast, poor decoders scored significantly lower than typical readers on the same measure in kindergarten only. However, poor decoders achieved significantly lower scores on the early measures of phonological awareness. Thus, the analyses suggest subgrouping was a valid methodological approach and that subgroups were fairly stable across time.

Using latent transition analysis (LTA), Catts et al. (2012) identified four latent classes of readers that were consistent across time, from kindergarten through tenth grade. The latent classes represented normal readers, readers with word reading disabilities, readers with comprehension disabilities, and readers with disabilities in word reading and comprehension. Results showed students moved between latent classes over time, which enabled the researchers to identify students with late-emerging reading comprehension difficulties; these students had a history of difficulty with linguistic comprehension. The authors posited that their reading comprehension difficulties may not have emerged until later grades because most early reading curricula place a heavier focus on decoding and word reading as compared to comprehension skills. The link between early oral language and reading comprehension skills has been documented elsewhere (Catts et al., 1999; Kendeou, van den Broek, et al., 2009; Nation et al., 2010; Storch & Whitehurst, 2002). Therefore, in an MTSS context, identifying early struggling readers and providing targeted intervention based on phonological and word reading skills may not be sufficient to ameliorate later reading comprehension difficulties for students with late-emerging reading difficulties. Additionally, identification schemes based solely on phonological and word reading skills will likely miss students with average word reading skills, but below average linguistic comprehension.

Compton, Fuchs, Fuchs, Elleman, and Gilbert (2008) used LTA to derive latent classes of readers from first through fourth grade. The authors identified two latent classes at first, second, and fourth grades, typically developing readers and students with reading disabilities. Results showed the latent classes were stable over time and the majority of students remained in their respective latent class across grades, but 7% of typically developing readers transitioned into the reading disabled class. This was evidence for late-emerging reading disabilities, which was an important finding as these students are difficult to detect in early elementary by definition. As these students are less likely to be identified as reading disabled in the early grades, they will not receive the necessary intervention provided in MTSS frameworks. One of Compton et al.'s (2008) aims was to explore the efficacy of the latent class indicators in terms of predicting late-emerging RD. However, the indicators were limited to two measures of word reading (word identification and sight word efficiency) and one measure of passage comprehension. The authors compared models with and without sight word efficiency and found that its inclusion reduced the number of false negatives; students who would typically not be identified as struggling readers, but do exhibit RD in later elementary. Importantly, these authors also identified linguistic comprehension, phonemic awareness, and phonemic decoding efficiency as promising first-grade predictors of late-emerging RD. While all three predictors produced false positives, linguistic comprehension produced the least (35) compared to phonemic awareness (79) and phonemic decoding efficiency (83). While the authors considered the number of false positives identified by linguistic comprehension to be unacceptable, it is worth noting linguistic comprehension was measured by a single variable. Using other or multiple measures of linguistic comprehension may have produced different results. Finally, the authors acknowledged measures other than early word reading need further exploration as early indicators of later RD. Given these findings and others that have shown students with late-emerging RD may have typically developing word reading skills (Badian, 1999; Chall, 1983; Leach, Scarborough, & Rescorla, 2003), it is reasonable to explore additional and multiple measures of reading sub-skills, such as linguistic comprehension, as indicators of RD.

Use of Latent Profile Analysis in the Present Study

Francis et al. (2005) argued that the use of cutoff scores on single variables can lead to instability in classification as children who score near the cutoff boundary may fluctuate between subgroups. Research methodologies that simultaneously examine multiple measures of reading skills to delineate reading profiles may produce more robust and accurate results. Very few studies have utilized mixture models to empirically identify latent subgroups of readers (i.e., not rely on predetermined cutoff scores), which is the approach used in the present study. An advantage of this approach is that individuals can be classified into subgroups according to patterns of scores across multiple variables. Additionally, classification criteria are data-driven rather than imposed by the researchers. While mixture models also suffer from classification error, this can be measured and accounted for by a given individual's posterior probability of belonging to a given latent profile. This provides researchers with a measure that can be used to judge the adequacy of the model, and this classification error can be accounted for when examining covariates, distal outcomes, or other auxiliary variables specified to be related to latent profile membership. In this study, we utilized multiple measures of each sub-skill to ensure the robustness of the emergent profiles (Francis et al., 2005). Finally, these analyses are model-based, which allows for replication. That is, models shown to be preferred in one study can be applied to independent samples and tested for their viability or rejection.

Latent profile analysis (LPA) is a probabilistic technique that is able to empirically derive categorically distinct profiles of individuals. This approach is similar to latent class analysis (LCA; which uses binary or ordinal variables), but is conducted with continuous variables (Gibson, 1959; Lazarsfeld & Henry, 1969). Figure 1 presents a conceptual diagram of the LPA in the current study. Separate LPAs were conducted for the subsamples of at-risk and not at-risk students, but we present one diagram, which is the same model for each subsample. The categorical latent profile variable is represented by a circle. Boxes above the circle represent each of the observed variables used to measure the latent profile variable, and the arrows indicate the latent variable is measured by the observed variable. The boxes to the right of the latent variable represent observed reading comprehension measures. Profile-specific means of these variables were estimated (represented by the arrows), but were not used to measure the latent profile variable. A separate procedure (described in the Data Analysis section) was used to estimate the profile-specific means of the reading comprehension variables.

Using multivariate data, individuals are classified together within a given profile because their pattern of responses across multiple measures are more similar to each other than individuals assigned to a different profile. Latent profiles differ in their means across variables, but can also differ in variances and covariances (Masyn, 2013). In this study, variances and covariances were constrained to equality across classes, which is commonly done for parsimony and is the default specification in M*plus*. Deriving profiles from responses - and simultaneously using the same responses to assign individuals to profiles - is a process that requires many iterations (i.e., using multiple random start values) as well as optimizing the resulting solutions. Additionally, the enumeration process used to choose the optimal model is iterative and relies on a number of considerations. Enumeration can be thought of as analogous to exploratory factor analysis and choosing the number of factors to retain based on multiple fit statistics and substantive interpretability. The researcher must specify the number of profiles to be estimated, collect and compare multiple fit statistics for all models, and check each model's interpretability. In this study, 500 random starts with 100 optimizations were used and models with 1 - 5 profiles were specified because there were model convergence problems with more than five profiles.

Present Study

In the present study, we use an independent sample to address multiple objectives cited in the individual studies mentioned above. The research questions for the present study are: 1) How many reader profiles would emerge for each risk status group and what would be the qualitative differences between the profiles?, 2) Which measures best differentiate the reading profiles?, and 3) How do the latent profiles differ in terms of reading comprehension achievement?

Methods

Participants

Data for this study come from a larger study investigating the efficacy of a reading intervention. All data in this study were collected prior to the beginning of the intervention. First-grade students were drawn from 30 classrooms and 15 schools from a suburban and rural region of California and an urban region of Texas. The sample consisted of N = 290 students, with 47.1% female and 77.7% qualifying for free lunch. The ethnic breakdown of the sample was 44.8% Latino(a), 23.4% African-American, 19.2% Caucasian, 3.5% Asian, and 8.7% “Other” or “Mixed.” Initial screenings (described below) identified n = 141 students as at-risk and n = 149 students as not at-risk. Nine or ten students were recruited from each classroom with approximately half of the recruited students considered at-risk. All students were enrolled in general education classrooms full time and none of the students were receiving special education services.

Determining At-Risk Status

At-risk status was determined using a two-step process. Risk status was used to identify students in need of the reading intervention provided by the larger project. First, teachers ranked all of their students according to their judgment of students’ reading skills, but they were not asked to officially designate risk status. Students ranked in the lower performing 50% of each class were then screened to identify those most at-risk of experiencing reading difficulties. Thus, while teacher judgment was used to identify potential students for screening, their judgment was not used to define risk status, nor was it a variable in this study. We used the Texas Primary Reading Inventory (TPRI; Children’s Learning Institute & Texas Institute for Measurement, Evaluation, and Statistics, 2010), which is designed to identify students in grades K-3 at-risk for reading difficulties and as a diagnostic tool. First, students were administered three screens assessing letter sounds, word reading, and blending phonemes. If students passed all three screens, they were considered not at-risk and were discontinued from the screening process. If students failed at least one of the three screens, they were administered a brief word-reading screen that consisted of simple words that students receive exposure to prior to the beginning of first grade. Prior research (Denton et al., 2010; Mathes et al., 2005) has demonstrated that this measure can successfully identify at-risk first graders. Finally, students were administered a listening comprehension screen from the TPRI. Examiners orally read two passages to students. Following each passage, examiners asked them to answer six literal and inferential questions, yielding a total of 12 questions. Students were considered at-risk if they scored five or below (out of 15) on the word-reading screen and had an average score of three or below (out of six) on the listening comprehension screen. These cut scores were chosen because they were consistent with prior studies (Denton et al., 2010; Mathes et al., 2005) that have used the TPRI. Students could only qualify as at-risk if they qualified in both word reading and listening comprehension constructs; students were disqualified if they exhibited difficulties in only one construct. We did not include classrooms with less than four students who qualified as at-risk. If a classroom contained more than five students who qualified, we selected the five lowest-performing students for purposes of the reading intervention. The not at-risk group was comprised of five students from each classroom who were randomly selected from the 50% of students who were teacher-nominated as not at-risk. Figure 2 presents a flow chart summarizing the selection of the at-risk and not at-risk subsamples.

Procedures

Examiners received group training on the assessment battery and were required to meet reliability criteria for each assessment with the trainers. Trainers administered the battery to the individual examiners in a role-playing situation. Examiners were required to administer each assessment with a minimum of 90% accuracy per assessment. Students were administered the assessment battery in quiet areas outside of their classrooms. To ensure the examiners continued to accurately administer the assessment battery, trainers checked all assessments at the end of each student's assessment session. Any errors were immediately corrected before the student returned to class. Data were collected within a 4-week period during September and October. Finally, all data were double-entered and, if errors occurred, were corrected by checking the original testing forms.

Measures

We converted scaled and standard scores to z-scores using national norms for all measures except QRI-5, which is non-normed. Scores on this measure were converted to z-scores using within-sample means and standard deviations.

At-Risk Screen. The TPRI was utilized to screen students to determine if they were at-risk for reading difficulties. The TPRI is designed to identify students in grades K-3 at-risk for reading difficulties and as a diagnostic instrument to assist in planning instruction. To screen first graders for risk, the following measures from the TPRI were utilized: Letter Sound Identification, Word Reading, Blending Phonemes, and Listening Comprehension. In the Letter Sound Identification subtest, students are presented with a short list of letters and asked to identify the most common sound for each item; the subtest has 10 items, reliability is reported at .91. The Word Reading measure contains 5 items and requires students to read the words out loud, reliability is .96. The Blending Phonemes subtest has 5 items, reliability is reported to be .92. For this subtest, the assessor presents individual phonemes to the student and they are required to use the phonemes presented to form a real word. The Listening Comprehension subtest requires students to listen to two passages and answer a series of questions orally after listening, with a total possible score on each passage of 6 points; reliability is reported at .72.

Phonological awareness. The Elision and Blending Words subtests of the Comprehensive Test of Phonological Processing (CTOPP; Torgesen, Wagener & Rashotte, 1999) were used to measure phonological awareness. For the Elision subtest, the examiner asks an individual to repeat a word presented orally while omitting one of the word's sounds (e.g., "say bold. Now say bold without /b/"). For the Blending Words subtest, an examiner plays a recording that asks the individual to combine separate sounds into whole words. For example, the first practice item presents the individual with "What word do these sounds make? Can-dy" with the correct response being candy. The examiner's manual reports a test-retest reliability of .88 for both subtests with individuals aged five to seven years. For the current sample, this study calculated Cronbach's alpha of .90 for Elision and .88 for Blending Words. In terms of validity, the CTOPP manual reports appropriate content-description, criterion-prediction, and construct-identification validity for these subtests. Additionally, both subtests have high factor loadings on a phonological awareness latent factor.

Decoding. Letter-Word Identification (LWID) and Word Attack (WA) subtests of the Woodcock-Johnson IV (WJ-IV; Schrank, Mather & Jaffe, 2014) were used to assess decoding skills. The LWID subtest presents individual letters/letter combinations and words in a list format. The WA subtest presents lists of pseudowords to assess phonological decoding. The technical manual reports split-half reliability coefficients for the ages in this study between .96 and .98 for LWID and .94 and .96 for WA. Cronbach's alpha for LWID for the current sample was .96 and .89 for WA.

Linguistic comprehension. Two separate measures of linguistic comprehension were utilized. First, the Understanding Spoken Paragraphs subtest from the Clinical Evaluation of Language Fundamentals, Fourth Edition (CELF-4; Semel, Wiig, & Secord, 2006) was used. An assessor read three short passages to students and asked five comprehension (literal and inferential) questions following each passage, yielding a total of 15 questions. Possible scaled scores range from 1 - 19. The test manual reports a Cronbach's alpha of .69 and .65 and split-half reliability of .74 and .73 for six-year olds. The test manual also reports multiple forms of validity evidence based on response process, structural equation models, and correlations with the CELF, Third Edition. Additionally, the manual reports validity with regard to varying demographic characteristics as well as varying disorders related to language development. Cronbach's alpha for the current sample was calculated as .81.

The second measure was an adapted version of the Qualitative Reading Inventory-5 (QRI-5; Leslie & Caldwell, 2011). Students were read a passage then asked six questions regarding explicit and implicit details of the passage. We analyzed the raw score (i.e., total number of correct responses) as this assessment is non-normed. We calculated a Cronbach's alpha of .66 for the study sample. The technical manual reports an alternate-form reliability, meaning instructional levels associated with two alternative stories matched 91% of the time for first graders. It also reports standard errors of measurement ranging from .14 - .16 for the stories used in this study. Cronbach's alpha for this sample was .65. The manual provides evidence of content, criterion-related, and construct validity.

Oral reading fluency. The fluency measures used was the Rate subscale of the Gray Oral Reading Test-5 (GORT-5; Wiederholt & Bryant, 2012). Students were asked to read aloud at least two passages. Students receive a score for each passage based on the number of seconds it took them to read the passage. We analyzed a final normed score that accounts for the total number of passages read. The examiner's manual reports Cronbach's alpha of .86 for six-year olds using Form A. Cronbach's alpha for this sample was .84. The technical manual also reports appropriate content-description, criterion-prediction, and construct-identification validity.

Reading comprehension. Two measures were used to assess reading comprehension. The first was the Passage Comprehension (PC) subtest of the WJ-IV, which requires individuals to read short sentences and passages with a missing word. They are asked to supply the word. The examiner's manual reports split-half reliability coefficients for this age range between .93 and .98. Cronbach's alpha for the current sample was .90. The second measure was the Comprehension score from the GORT-5. Individuals read increasingly complex passages and were asked five open-ended comprehension questions after each passage. This test is discontinued when students are unable to fluently read two consecutive passages according to the test's stopping criterion. The technical manual reports Cronbach's alpha of .92 for this age range and was .72 for this sample.

Data Analysis Plan

First, we compared the teacher-nominated at-risk and not at-risk groups across all variables to ensure the groupings were supported empirically. We conducted a series of *t*-tests with all of the variables. To account for multiple comparisons, we employed a Bonferroni correction and used an adjusted *p*-value of .006 as the cutoff for statistical significance.

The latent profile analyses (LPA) were conducted in multiple steps with each subgroup separately. The same steps were performed for each subgroup. A conceptual diagram is presented in Figure 1. To account for the nested nature of the data, we clustered students at the teacher level. We began by fitting a model with one profile then increased the number of profiles by one with each subsequent iteration. Model fit statistics were recorded and compared to determine the optimal number of profiles within each subgroup. Additionally, we used substantive reasoning to ensure the chosen model made conceptual sense (Muthén, 2003). All models were run using M*plus* version 7.4 (Muthén & Muthén, 1998 – 2015) using Full Information Maximum Likelihood estimation. This estimator allowed students to be included as long as they had data on at least one observed variable unless they were missing data on the distal outcomes (i.e., reading comprehension variables). As these models are known to converge on local solutions, we used a large number of random start values (McLachlan & Peel, 2000).

We used commonly employed fit statistics, specifically the Bayesian Information Criterion (BIC), Adjusted BIC (ABIC) where lower values indicated better fit. We also used two likelihood-based indices, the Lo-Mendell-Rubin Likelihood Ratio Test (LMR) and the Bootstrap Likelihood Ratio Test (BLRT). These models provide a p-value to compare a k profile model to a k - 1 profile model. A non-significant p-value indicates the additional profile did not significantly improve the model. Finally, we also calculated values for the Approximate Weight of Evidence (AWE) criterion, with lower values indicating better fit. See Nylund, Asparouhov, and Muthén (2007) and Masyn (2013) for additional information on the fit statistics.

Once the preferred unconditional model (i.e., without distal outcomes) was chosen, the distal outcomes (i.e., reading comprehension variables) were included using the BCH approach (Asparouhov & Muthén, 2014; Bakk & Vermunt, 2014; Bolck, Croon, & Hagenaars, 2004; Vermunt, 2010). This approach avoids shifts in the latent profiles that can occur when auxiliary variables (i.e., distal outcomes) are included in the model. After choosing the preferred unconditional model, weights are applied to each individual based on posterior probabilities of membership in each latent profile. This treats latent profile membership as known. Finally, the distal outcome variables are included and a multiple group analysis is performed treating the latent profiles as observed subgroups. This approach avoids shifts in the latent profiles when the distal outcomes are included, which can occur with other methods of including distal outcomes in mixture models (Asparouhov & Muthén, 2014). Mean estimates of the two reading comprehension variables were then estimated for each latent profile and compared for significant differences. Within each risk group, significant differences were examined using a series of Wald tests, which were provided by the BCH approach and available in the M*plus* output. To test for significant differences across risk groups, we calculated pooled variances and conducted *t*-tests. This yielded six *t*-tests for each reading comprehension measure. To control for familywise error rate, we applied a Bonferroni correction and used a *p*-value of .008 to indicate significant differences.

Results

Descriptive Statistics

Table 1 presents descriptive statistics disaggregated by risk status. The mean scores for the not at-risk group were significantly higher compared to the at-risk group across all variables, as indicated by *t* values in the far right column of Table 1. All *t*-tests were significant at $p < .001$. Additionally, the maximum score achieved by the not at-risk group was greater than the maximum score achieved by the at-risk group for every variable except QRI-5. For most variables, this was also true of the lowest score achieved by each subgroup.

Latent Profiles for At-risk Readers

We examined models consisting of one to four profiles as models with five and six profiles did not converge. Fit statistics can be seen in Table 2. Neither the BIC nor ABIC reached a minimum value. However, the AWE reached a minimum value at two profiles. Similarly, the LMR value approached significance at two profiles, but not at three or four profiles. The BLRT never became non-significant and was considered uninformative. Next, we examined the item-profile plots to ensure the emergent profiles were substantively meaningful (Muthén, 2003). The two-profile plot is discussed in further detail below. The three- and four-profile models contained profiles that were redundant with the two-profile model. Specifically, the three-profile model contained two profiles with overlapping means (within .3 z-score points) on four of the seven indicators. The four-profile model contained two profiles with overlapping means on six of the seven indicators and another two profiles with overlapping means on four of the seven indicators. Moreover, one of the redundant profiles in the three- and four-profile models consisted of a small proportion of the sample (5.7%), which was further evidence they were not theoretically viable. Thus, we retained the two-profile model.

Figure 3 presents the item-profile plot of all profiles. Though we conducted LPAs separately with at-risk and not at-risk readers, we present all profiles on a single plot to foster interpretability and comparisons. Both the separate and combined item-profile plots were used to interpret and the label the emergent plots. The two profiles of at-risk readers can be seen at the bottom Figure 3 and are denoted by solid lines. The profile at the very bottom of the plot scored lowest on all items and is demarcated by circular markers. They scored nearly two standard deviations below average on five of the seven measures. Thus, we labeled this profile *At-Risk Global* and they consisted of 21.0% of the overall sample. The average posterior probability for this profile was .94. The profile demarcated by a solid line with triangle markers performed below average on most measures, but was marked by poor performance on the GORT rate measures. This profile was labeled *At-Risk Fluency* and consisted of 27.6% of the overall sample and the average posterior probability was .95.

Latent Profiles for Not At-risk Readers

We examined models consisting of one to five profiles as the model with six profiles suffered from convergence problems. The two-profile model was supported by the AWE. The three-profile model was supported by the LMR *p*-value. However, findings with the BIC were more nuanced. A minimum BIC value supported the four-profile model, but it has also been shown that the point at which the BIC displays an “elbow” (i.e., subsequent BIC values indicate diminished returns) may also be indicative of the preferred model (Nylund et al., 2007). This elbow occurred with the three-profile model.

Having statistical support for the two-, three-, and four-profile models, we proceeded by examining these item-profile plots for substantive consideration (Muthén, 2003). The two-profile plot consisted of one profile that scored approximately one standard deviation above average on phonological awareness and decoding variables, but between average and one-half standard deviation above average on linguistic comprehension and fluency variables. The second profile scored between average and one standard deviation on all variables. The three-profile model was better differentiated in that one profile consistently scored above average, one profile consistently scored near average, and one profile consistently scored near or below average. The four-profile model included a profile that was redundant with at least one other profile on six of the seven measures. Thus, the four-profile model was not parsimonious and was considered substantively untenable. We retained the three-profile model as it had both statistical and substantive support.

The three profiles are presented in Figure 3 and are denoted by dashed lines. The profile with square markers scored remarkably similar across all measures to the *At-Risk Low Fluency* profile. Though their scores nearly matched the *At-Risk Low Fluency* profile, we termed them *Not At-Risk Low Fluency* to reflect teacher nomination of risk status. However, we emphasize the term "*Not At-Risk*" is based on teacher judgment and we elaborate on this further in the Discussion section. This profile was composed of 17.2% of the overall sample and the average posterior probability was .94. The profile demarcated by diamond markers hovered between one-half of a standard deviation below and above average on all measures. We labeled this profile *Average* and it consisted of 21.4% of the students and the average posterior probability was .92. Finally, the profile at the top of the plot denoted by X markers scored highest on all measures. This profile scored near or above one standard deviation on the phonological awareness and decoding measures as well as above average on linguistic comprehension and fluency measures. We termed this profile *Above Average* and it consisted of 12.8% of the overall sample with an average posterior probability of .97.

Differentiating Profiles by Measure

Comparing measures in terms of their ability to differentiate reading profiles can be done by visually examining the distance between profiles on each measure (see Figure 3). The phonological awareness and decoding measures exhibited the greatest variation across profiles. However, the *Not At-Risk Low Fluency* and *At-Risk Low Fluency* profiles exhibited nearly identical mean scores across these sets of measures. The remaining three profiles were well-differentiated by these measures.

The linguistic comprehension measures performed in parallel across profiles. That is, all profiles scored higher, on average, on the QRI-5 measure than the CELF measure. The *At-Risk Global* profile scored markedly lower on these two measures than the other four profiles. The *Above Average* profile was also clearly delineated from the other profiles by these two linguistic comprehension measures. However, the *Average*, *Not At-Risk Low Fluency*, and *At-Risk Low Fluency* profiles were not well-differentiated by CELF and QRI-5.

GORT rate delineated the three not at-risk profiles, but there was virtually no differentiation among the at-risk students.

Reading Comprehension Prediction

We included two measures of reading comprehension as distal outcomes of the latent profiles, which allowed us to estimate profile-specific means. Figures 4 and 5 summarize the results for WJ Passage Comprehension and GORT-5 Comprehension, respectively. Results were nearly identical for both reading comprehension outcomes, lending validity to the emergent profiles. In terms of WJ Passage Comprehension, *At-risk Global* had the lowest mean (M = 75.37), which was more than one and a half standard deviations below the national average. Both *At-risk Low Fluency* (M = 86.61) and *Not At-risk Low Fluency* (M = 88.63) were within one standard deviation below the national norm. The *Average* profile (M = 100.32) scored at the national average and the *Above Average* profile (M = 114.80) scored approximately one standard deviation above the national average. The means for the *At-Risk Low Fluency* and *Not At-Risk Low Fluency* profiles were not significantly different (p = .25). All other pairwise mean comparisons were significant at $p < .001$.

A slightly different picture emerged when examining the means of GORT-5 comprehension, though the rank ordering of the means remained the same. Both the *At-risk Global* (M = 3.96) and *At-risk Low Fluency* (M = 4.63) profiles scored approximately two standard deviations below the national norm. The *Not At-risk Low Fluency* profile (M = 5.96) scored more than one standard deviation below the national norm. This was in contrast to scores on the WJ Passage Comprehension measure in which both the *At-risk Low Fluency* and *Not At-risk Low Fluency* profiles scored within one standard deviation of the national average. The *Average* profile (M = 8.50) scored below the national average, but was within one standard deviation. The *Above Average* profile (M = 11.07) scored within one standard deviation above the national norm. All pairwise mean comparisons were significant at $p < .001$, including the comparison of the *At-Risk Low Fluency* and *Not At-Risk Low Fluency* means, which contrasted with the WJ Passage Comprehension findings.

Discussion

Heterogeneous Reading Profiles

The goal of this study was to examine whether meaningful subgroups of at-risk and not at-risk readers could be identified in a sample of first grade readers. Characterizing subgroups of readers using multiple measures may better identify students who require intervention and subsequently inform intervention design that accounts for students' nuanced skill profiles and better targets individual needs. This approach may also help educational practitioners identify students who hover near cutoffs for meeting at-risk criteria. An important feature of this study was identifying these subtypes empirically rather than using somewhat arbitrary cutoff scores to classify students, as has been used in previous research (Catts et al., 2003; Catts et al., 2006; Nation et al., 2010). Second, this study compared measures of multiple subcomponent skills of reading comprehension to examine which measures best differentiated the emergent subgroups. Finally, we explored whether the emergent subgroups were consequential in terms of reading comprehension achievement. Establishing this relation has important implications as it would provide evidence that students at-risk for reading comprehension difficulties require varied approaches to intervention. In order to answer the study's research questions, two latent profile analyses were conducted to identify subgroups of at-risk readers and subgroups of not at-risk readers based on subcomponent skills of reading comprehension. Two subgroups of at-risk readers and three subgroups of typically developing readers were identified. Next, two measures of reading comprehension as a function of subgroup membership were examined. Categorically distinct subgroups were identified; however, these subgroups were generally rank-ordered, especially with respect to the phonological awareness and decoding variables. While this finding can be interpreted as suggestive of a continuum of skill development, we argue that pinpointing an individual student's prereading skills along a continuum is less beneficial to applied educators than a heuristic that can group students with similar proficiencies. This is because, it is possible, that the latter approaches provide a more useful foundation to design intervention for multiple students simultaneously. We discuss these findings as well as implication for practices and limitations of the current study below.

A critical component of the MTSS framework is the identification of students who are at-risk for developing later reading difficulties. Early identification is important, as extant data demonstrate the need for intensive, early reading intervention for optimal outcomes (Francis et al., 1996; Lyon & Fletcher, 2001; Torgesen, Rashotte, & Alexander, 2001). This study demonstrates that early identification of at-risk readers is a difficult task. Teachers were asked to rank their students in terms of reading skills, and only students in the bottom half of the rankings were administered the screener to determine at-risk status. Yet, a surprising finding emerged, in which the *At-Risk Low Fluency* and *Not At-Risk Low Fluency* profiles bore a striking similarity in terms of achievement across the measures (see Figure 3). This suggests that the 17.2% of students who comprised the *Not At-Risk Low Fluency* profile would likely benefit from reading intervention, but they were not identified as such by their teachers in the beginning of first grade. That is, even though these students were not ranked by their teachers as performing in the lower 50% of their class, there are some students who hover near cutoffs for risk and, while difficult to identify, these students may also be candidates for some amount of intervention. This finding is consistent with prior research investigating teachers' judgments regarding students' oral reading fluency may not be accurate (Hamilton & Shinn, 2003; Meisinger et al., 2009). These findings support previous researchers who suggest that universal screening is the best way to identify at-risk readers in the early grades (Glover & Albers, 2007; Jenkins, Hudson, & Johnson, 2007).

In contrast to the two *Low Fluency* profiles, the subgroup that would likely be most easily identified as at-risk for reading difficulties in first grade was the *At-Risk Global* profile. Notably, this profile contained one-fifth of our study sample. While it may be expected that students in this profile would score markedly lower on measures of phonological awareness and word reading, this profile also exhibited substantial difficulties with linguistic comprehension. Thus, screening procedures might be improved by including measures of linguistic comprehension in the early grades. There is an emerging body of evidence that supports the relation between linguistic comprehension and later reading comprehension (Catts et al., 1999; Kendeou, van den Broek, et al., 2009; Ouellette & Beers, 2010; Storch & Whitehurst, 2002). Early reading intervention often focuses on word reading skills (Foorman et al., 1997; Foorman et al, 1998; Rashotte, MacPhee, & Torgesen, 2001); our results suggest targeting linguistic comprehension skills for the most at-risk students may also be necessary. Remediating word reading skills may result in short-term gains in reading fluency and reading comprehension, but students may continue to struggle (Denton et al., 2013) or students might develop reading difficulties later in school (Catts et al., 2001; Compton et al., 2008) as text becomes more complex. Linguistic comprehension skills have been shown to be a stronger predictor of reading comprehension in later grades (Hoover & Gough, 1990; Kershaw & Schatschneider, 2012; Storch & Whitehurst, 2002; Vellutino et al., 2007), so early interventions that include both word reading and linguistic comprehension may be more beneficial than interventions focused solely on word reading skills with respect to long-term reading comprehension outcomes.

Most Efficient Measures to Differentiate Reading Profiles

Given our results concerning the importance of linguistic comprehension, is reading fluency a sufficient measure to be used unilaterally to predict reading comprehension performance? The GORT rate measure differentiated the three profiles of not at-risk students, but did not perform well in terms of delineating the students most at-risk. The *At-Risk Global*, *At-Risk Low Fluency*, and *Not At-Risk Low Fluency* subgroups performed similarly with respect to reading fluency. However, the *At-Risk Global* profile was most differentiated from the other two profiles by the other subcomponents of reading comprehension: phonological awareness, decoding, and linguistic comprehension. Therefore, in first grade, these three skill domains appear better able to distinguish students who present the greatest risk for reading comprehension difficulties. This is not to suggest that measuring reading fluency in first grade is not beneficial. The reading fluency measure did reliably distinguish the two highest performing profiles (*Average* and *Above Average*) from the remaining profiles, but identifying such coarse differences may be of limited utility when the focus is on identifying students in need of intervention. Had reading fluency been used as a unilateral measure to differentiate students, it is unlikely that it would have distinguished between the three lower performing profiles. With respect to predicting reading comprehension difficulties, reading fluency may be useful, but other domains, including linguistic comprehension should also be considered.

The measures with the greatest variation and, thus, the greatest potential to differentiate first grade readers were the phonological awareness and decoding measures. This is consistent with prior empirical research that has demonstrated the link between these skills and reading comprehension in early elementary (Chen & Vellutino, 1997; Georgiou et al., 2009; Hoover & Gough, 1990; Joshi & Aaron, 2000). However, as noted above, there was a surprising similarity between the *At-Risk Low Fluency* and *Not At-Risk Low Fluency* profiles on all phonological awareness and decoding variables. Even though the phonological awareness and decoding measures demonstrated the most potential to differentiate students, these measures essentially were unable to delineate these profiles. Together, these profiles accounted for approximately 45% of the sample. This is a sizeable number of students who appear to be performing slightly below average on most of the phonological awareness and decoding measures. As such, these students may be thought to be on the cusp for experiencing future reading difficulties. This suggests that students who perform below average, but perhaps not overtly so, on phonological awareness and decoding measures should be referred for screening if universal screening is not available. It is possible that both profiles of students would benefit from Tier 2 support that would alleviate future reading difficulties.

The CELF and QRI-5 were used to measure linguistic comprehension. Overall, there was less variation in these measures across the profiles compared to both the phonological awareness and decoding measures, but this is consistent with the Simple View with respect to early elementary students (e.g., Hoover & Gough, 1990; Kendeou, van den Broek, et al., 2009; Vellutino et al., 2007). Thus, it was not surprising these two measures did not differentiate the three middle profiles in Figure 3. However, both measures clearly delineated the *At-Risk Global* profile from the remaining profiles. This suggests these linguistic comprehension measures may be used in conjunction with the phonological awareness and decoding measures when the aim is to identify students with the greatest need for intervention.

The CELF and QRI-5 performed in parallel across the profiles. That is, every profile scored higher on the QRI-5 than the CELF. This was somewhat surprising as the CELF is a normed measure whereas the QRI-5 is not. Since these results were consistent across profiles, these measures may be considered to be providing similar information regarding students' linguistic comprehension skills. This is a promising finding as the QRI-5 is a commonly used assessment by teachers and school psychologists. A drawback, however, is that administering a single early elementary QRI-5 story - as was done in this study - yields six comprehension questions compared to administering three passages from the CELF Understanding Spoken Paragraphs subtest, which contains 15 comprehension questions. Fewer questions may make it more difficult for teachers and school psychologists to interpret QRI-5 results as there is an increased potential for at-risk students' scores to be artificially inflated by capitalizing on chance. This may explain why both the at-risk and not at-risk subgroups had the same range of scores (see Table 1) on the QRI-5. While the range of scores on the CELF was similar for both subgroups, there was greater variation in the means of each subgroup compared to the QRI-5 means. Moreover, since the CELF is normed, results are more readily interpreted and may be more easily communicated to parents and other education professionals.

Predicting Reading Comprehension Achievement

Profile membership was used to predict two measures of reading comprehension, the passage comprehension subtest of the WJ-IV and the comprehension subscale of the GORT. Generally, results were as might be expected intuitively. The *Above Average* profile, which performed the highest on all measures, also achieved the highest means on both reading comprehension measures, while the *At-Risk Global* profile, which performed lowest on all measures, achieved the lowest reading comprehension means. The reading comprehension means of the remaining profiles were also aligned with the general rank ordering of achievement patterns across the measures as depicted in Figure 3. These intuitive findings lend validity to the emergent profiles and support the notion of categorically distinct latent profiles as a useful heuristic for school practitioners seeking to identify readers in need of intervention.

The reading comprehension results showed significant differences for the vast majority of the pairwise mean comparisons between profiles. All of the profile-specific means of the GORT comprehension subscale were significantly different from each other. However, the WJ-IV passage comprehension means of the *At-Risk Low Fluency* and *Not At-Risk Low Fluency* profiles were not significantly different from each other. Additionally, the WJ-IV passage comprehension means of these two profiles were both within one standard deviation of the national norm, while the GORT comprehension means of these two profiles were both more than two standard deviations below the national norm. This may reflect differences between the formats of the two measures. The WJ-IV passage comprehension subtest presents a rebus and cloze format while the GORT comprehension subtest utilizes open-ended questions read by the examiner. Additionally, the stopping criterion for the GORT is met when a student is unable to read the passage fluently, whereas the stopping criterion for the WJ-IV passage comprehension is met when a student misses six consecutive items. Differences in format may indicate these two reading comprehension tasks rely on reading sub-skills in different ways. Keenan, Betjemann, and Olson (2008) examined prior versions of these subtests and found only a modest correlation. Moreover, their study demonstrated the WJ passage comprehension subtest was more strongly related to decoding than the GORT, and suggested this may be because the WJ consists of shorter passages. This is also true of the newer versions utilized in this study, so their rationale may also explain the differences found here. Finally, and perhaps, more importantly, Keenan et al. (2008) found the relation between decoding and reading comprehension for the WJ passage comprehension subtest was dependent on the level of reading development with the effect being stronger for younger and less skilled readers. Therefore, the non-significant difference in the WJ-IV passage comprehension means between the *At-Risk Fluency* and *Not At-Risk Low Fluency* profiles may reflect the fact that their decoding skills were strikingly similar. Keenan et al. (2008) found the GORT reading comprehension was dependent on both decoding and linguistic comprehension. In the present study, the significant differences between the *At-Risk Fluency* and *Not At-Risk Low Fluency* profiles on GORT reading comprehension may simply stem from the fact that multiple measures across multiple reading sub-skills were utilized, which could have enhanced the profile-specific differences on the GORT.

These results may also be related to the difficulty in identifying at-risk students who hover near the middle range of reading achievement. As discussed earlier, 17.2% of our sample (the *Not At-Risk Low Fluency* profile) was deemed not at-risk by teachers, though our analyses demonstrate their reading skill profile was almost identical to a subgroup of 27.6% of our sample the teachers deemed at-risk (the *At-Risk Low Fluency* profile). This speaks to the importance of universal screening to identify students who appear to be on the cusp of requiring reading intervention. Further, Tier 2 intervention might provide a greater benefit to students on this cusp than general curricula since at-risk status is difficult to discern. If educators wish to be cautious, then it would behoove them to refer all of the middle-performing students for further screening.

Implications for Practice

The results of this study suggest screening procedures used in first grade should begin with phonological awareness and decoding measures as these exhibited the greatest variation. This is consistent with literature exploring the utility of gated screening with this age group

(Clemens, Shapiro, & Thoemmes, 2011; Compton et al., 2010). Gated screening involves first administering an efficient, brief assessment to all students followed by a more comprehensive battery of assessments to those who may be at-risk. Results from this study suggest phonological awareness and decoding assessments are the first screenings that should be conducted. Since the *At-Risk Low Fluency* profile demonstrated concomitant weaknesses in linguistic comprehension, these assessments should be included in a secondary battery. Since the *At-Risk Low Fluency* profile exhibited weaknesses in both areas, it would seem phonological awareness and decoding assessments would be sufficient to identify them. However, extant literature has identified profiles of readers who demonstrated adequate decoding skills, but experienced difficulty with linguistic comprehension (e.g., Catts et al., 2006; Nation et al., 2010). Though this study was unable to identify a comparable group of students, prior findings would suggest best practice would include a screening measure of linguistic comprehension in addition to decoding. For instance, Compton et al. (2010) included a measure of oral vocabulary in their gated screening procedures and found it was informative. Restricting interventions to word reading skills leaves linguistic comprehension skills - a critical subcomponent of reading comprehension in later grades - unaddressed, which may lead to the re-emergence of reading comprehension difficulties as students progress through school.

Teachers often assess oral reading fluency as an indicator of reading progress. This study supports this practice, but only in terms of examining coarse differences among students. That is, the reading fluency measure used in this study delineated the two highest performing profiles from the three lower performing profiles. It did a much poorer job of distinguishing among students in the three lower performing profiles, especially when compared to the other measures used in this study. It is likely that these students' word reading skills were not developed enough to allow for much variation in their oral reading fluency skills. Thus, teachers and other practitioners may be measuring oral reading fluency before it is able to provide useful information. For instance, Silverman, Speece, Harring, and Ritchey (2013) found oral reading fluency mediated the relation between decoding and reading comprehension, but in a sample of fourth grade students. Moreover, prior studies of first grade reading screening have shown fluency in identifying individual words is a more effective measure than fluency in reading connected text when identifying first grade struggling readers (Clemens et al., 2011; Compton et al., 2010). This does not suggest oral reading fluency should not be a target of early intervention; indeed, we advocate for its inclusion as a component of reading intervention. However, basing instructional and intervention decisions on oral reading fluency may be fruitless, or worse, misguided with respect to at-risk first graders. A better strategy would be to tailor intervention based on word reading and linguistic comprehension skills until word reading enables oral reading fluency to reach a proficiency level that is informative.

Limitations

LPA is generally regarded as a large-sample technique and the size of each subsample in this study borders on the lower end of what is usually considered acceptable. However, given that there were only a small number of profiles identified within each subsample, and that each profile had a high average posterior probability (all >.92), there is evidence of a strong signal in the data supporting these findings. Nevertheless, future studies should examine whether these findings can be replicated in other samples of early first grade students, accounting for demographic characteristics. Another limitation attributable to sample size is this study did not identify a profile of readers who exhibited adequate decoding skills, but poor linguistic comprehension, which has been found in prior studies (Catts et al., 2006; Nation et al., 2010). Larger numbers of each subsample may have allowed this type of profile to be identified.

Second, accurately measuring reading comprehension in first grade is a difficult task, as evidenced by the inconsistent results concerning the WJ and GORT comprehension measures. While the rank ordering of profile-specific reading comprehension scores aligned was consistent, statistically significant differences among the profiles were not. Thus, these results may be more informative for the design of interventions targeting reading sub-skills than reading comprehension directly.

Third, it is unclear how the results might have changed if a different sampling procedure was used. For example, the not at-risk sample did not include students who were ranked in the lower 50% of each classroom, but passed the screening. If these students were re-included in the not at-risk sample, the nature and proportions of the not at-risk profiles might have been different. Perhaps the *Not At-Risk Low Fluency* might have contained a larger proportion of the not at-risk students. Additionally, screening all students might have yielded a different at-risk sample compared to screening only students ranked by teachers in the lower 50% of their classrooms in terms of reading skills. Furthermore, students were only included in the at-risk sample if they failed all of the three screening measures. It is possible that a more liberal benchmark for defining risk - such as including all students who did not pass any one of the first three screening measures - could have resulted in a different at-risk sample, which could have yielded different profiles. While it is impossible to know exactly how these decisions might have impacted the results, it is likely that the changes listed above would have mostly affected the *At-Risk Low Fluency* and *Not At-Risk Low Fluency* profiles as those profiles exhibited the greatest overlap.

Conclusions

Prior research has often defined risk status by utilizing relatively arbitrary cutoff scores on individual measures or combinations of them. Recent advances in statistical modeling now enable researchers to empirically classify students based on multivariate response patterns. This study identified distinct profiles of at-risk and not at-risk students in the fall of first grade. However, this study was cross-sectional and it is not clear if these profiles would emerge in other grades or even at later points in first grade. Furthermore, while relations between the emergent profiles and reading comprehension were apparent, we do not know if these relations would be stable over time, especially if the latent profiles themselves are not stable. Further research in these areas is warranted.

Identifying at-risk readers in early elementary is a difficult task, and this study suggests teacher judgement is not sufficient. This supports universal screening, but the measures used in these procedures should be carefully examined for their utility. We found oral reading fluency to be least useful with this sample, but this may not apply to other samples, ages, or grades. Linguistic comprehension was identified as a critical risk factor for reading difficulties and this should be a target of intervention in early elementary. Little is known about the long-term effects of early linguistic comprehension on later reading comprehension, but this is a promising area of study that deserves further attention and may enable educational practitioners to develop better reading interventions.

References

Adlof, S.M., Catts, H.W., & Lee, J. (2010). Kindergarten predictors of second versus eighth grade reading comprehension impairments. *Journal of Learning Disabilities, 43*, 332-345. doi: 10.1177/0022219410369067

Adlof, S.M., Catts, H.W., & Little, T.D. (2006). Should the simple view of reading include a fluency component? *Reading and Writing, 19*, 933-958. doi: 10.1007/s11145-006

Asparouhov, T., & Muthén, B. (2014b). *Auxiliary variables in mixture modeling: Using the BCH method in Mplus to estimate a distal outcome model and an arbitrary secondary model* (Mplus Web Note: No. 21). Retrieved from: http://www.statmodel.com/download/asparouhov_muthen_2014.pdf

Badian, N. A. (1999). Reading disability defined as a discrepancy between listening and reading comprehension. *Journal of Learning Disabilities, 32*, 138-148. doi:10.1177/002221949903200204

Bakk, Z. & Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling: An Interdisciplinary Journal, 23*(1), 20-31. doi: 10.1080/10705511.2014.955104

Begeny, J.C., Eckert, T.L., Montarello, S.A., & Storie, M.S. (2008). Teachers' perceptions of students' reading abilities: An examination of the relationship between teachers' judgments and students' performance across a continuum of rating methods. *School Psychology Quarterly, 23*, 43-55. doi: 10.1037/1045-3830.23.1.43

Bolck, A., Croon, M., & Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis, 12*, 3-27. doi: 10.1093/pan/mph001

Bowyer-Crane, C., Snowling, M.J., Duff, F.J., Fieldsend, E., Carroll, J.M., Miles, J., …Hulme, C. (2008). Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology and Psychiatry, 49*, 422-432. doi: 10.1111/j.1469-7610.2007.01849.x

Cain, K., & Oakhill, J. (2011). Matthew effects in young readers: Reading comprehension and reading experience aid vocabulary development. *Journal of Learning Disabilities*, *44*, 431-443. doi: 10.1177/002221941140042

Catts, H.W., Adlof, S.M., Weismer, S.E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49,* 278-293. doi: 10.1044/1092-4388(2006/023)

Catts, H.W., Compton, D., Tomblin, J.B., & Bridges, M.S. (2012). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology, 104*, 166-181. doi: 10.1037/a0025323

Catts, H.W., Fey, M.E., Zhang, X., & Tomblin, B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading, 3*, 331-361. doi: 0.1207/s1532799xssr0304_2

Catts, H.W., Hogan, T.P., & Fey, M.E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities, 36,* 151-164. doi: 10.1177/002221940303600208

Chall, J. S. (1983). *Learning to read: The great debate*. New York: McGraw-Hill.

Chard, D. J., Vaughn, S., & Tyler, B. J. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities*, *35*(5), 386-406. doi: 10.1177/00222194020350050101

Chen, R.S. & Velluntino, F.R. (1997). Prediction of reading ability: A cross-validation study of the simple view of reading. *Journal of Literacy Research*, *29*, 1-24. doi: 10.1080/10862969709547947

Clemens, N.H., Shapiro, E.S., & Thoemmes, F. (2011). Improving the efficacy of first grade reading screening: An investigation of word identification fluency with other early literacy indicators. *School Psychology Quarterly*, *26*, 231-244. doi: 10.1037/a0025173

Compton, D.L., Fuchs, D., Fuchs, L.S., Elleman, A.M., & Gilbert, J.K. (2008). Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. *Learning and Individual Differences, 18*, 329-337. doi: 10.1016/j.lindif.2008.04.003

Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Gilbert, J. K., Barquero, L. A., ... &

Crouch, R. C. (2010). Selecting at-risk first-grade readers for early intervention: Eliminating false positives and exploring the promise of a two-stage gated screening process. *Journal of Educational Psychology*, *102*, 327-340. doi: 10.1037/a0018448.

Denton, C. A., Nimon, K., Mathes, P. G., Swanson, E. A., Kethley, C., Kurz, T. B., & Shih, M. (2010). Effectiveness of a supplemental early reading intervention scaled up in multiple schools. *Exceptional Children*, *76*(4), 394-416. doi: 10.1177/001440291007600402

Denton, C.A., Tolar, T.D., Fletcher, J.M., Barth, A.E., Vaughn, S., & Francis, D.J. (2013). Effects of tier 3 intervention for students with persistent reading difficulties and characteristics of inadequate responders. *Journal of Educational Psychology, 105*, 633-648. doi: 10.1037/a0032581

Fletcher, J.M. & Vaughn, S. (2009). Response to intervention: Preventing and remediating academic difficulties. *Child Development Perspectives, 3,* 30-37.

doi: 10.1111/j.1750-8606.2008.00072.x

Foorman, B.R., Francis, D.J., Fletcher, J.M., Schatschneider, C., Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*, 37-55. doi: 10.1037//0022-0663.90.1.37

Foorman, B.R., Francis, D.J., Shaywitz, S.E., Shaywitz, B.A., & Fletcher, J.M. (1997). The case for early reading intervention. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 243-264). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Francis, D. J., Fletcher, J. M., Stuebing, K. K., Lyon, G. R., Shaywitz, B. A., & Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities*, *38,* 98-108. doi:10.1177/00222194050380020101

Francis, D.J., Shaywitz, S.E., Stuebing, K.K., Shaywitz, B.A., & Fletcher, J.M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology*, 88, 3-17. doi: 10.1037//0022-0663.88.1.3

Fuchs, D. & Fuchs, L.S. (2006). Introduction to response to intervention: What, why, and how valid is it? R*eading Research Quarterly*, *41,* 93-99. doi:10.1598/rrq.41.1.4

Fuchs, L.S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. Scientific Studies of Reading, 5, 239–256. doi: 10.1207/s1532799xssr0503_3

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal measures of reading comprehension. Remedial and Special Education, 9, 20–28.

doi: 10.1177/074193258800900206

Fuchs, D., Mock, D., Morgan, P.L., & Young, C.L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice, 18,* 157-171. doi: 10.1111/1540-5826.00072

Georgiou, G.K., Das, J.P., & Hayward, D. (2009). Revisiting the "Simple View of Reading" in a group of children with poor reading comprehension. *Journal of Learning Disabilities, 42*, 76-84. doi: 10.1177/0022219408326210

Gerber, M.M. (2005). Teachers are still the test: Limitations of response to instruction strategies for identifying children with learning disabilities. *Journal of Learning Disabilities, 38,* 516-524. doi: 10.1177/00222194050380060701

Gibson, W.A. (1959). Three multivariate models: Factor analysis, latent structure analysis, and latent profile analysis. *Psychometrika*, *24,* 229-252.

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*, 117-135. doi: 10.1016/j.jsp.2006.05.005

Gough P.B. & Tunmer, W.E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7, 6-10. doi: 10.1177/074193258600700104

Hamilton. C., & Shinn, M. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review*, 32, 228-240.

Hoover, W.A., & Gough, P.B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2,* 127-160. doi: 10.1007/bf00401799

Jenkins, J.R., Fuchs, L.S., van den Broek, P., Espin, C., & Deno, S.L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of*

Educational Psychology, 95, 719-729. doi: 10.1037/0022-0663.95.4.719

Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for at-risk readers in a response to intervention framework. *School Psychology Review*, *36*, 582-600.

Joshi, R.M. & Aaron, P.G. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology, 21,* 85-97. doi: 10.1080/02702710050084428

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading, 12,* 281-300. doi: 10.1080/10888430802132279

Kendeou, P., Savage, R., & van den Broek, P. (2009). Revisiting the simple view of reading. *British Journal of Educational Psychology, 79,* 353-370. doi: 10.1348/978185408X369020

Kendeou, P., van den Broek, P., White, M.J., & Lynch, J.S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101,* 765-778. doi: 10.1037/a0015956

Kershaw, S. & Schatschneider, C. (2012). A latent variable approach to the simple view of reading. *Reading and Writing, 25,* 433-464. doi: 10.1007/s11145-010-9278-3

Kim, Y.S., Petscher, Y., Schatschneider, C., Foorman, B. (2010). Does growth rate in oral reading fluency matter in predicting reading comprehension achievement? *Journal of Educational Psychology, 102*, 652-667. doi: 10.1037/a0019643

LaBerge, D., & Samuels, S. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293–323. doi: 10.1016/0010-0285(74)90015-2

Lazarsfeld, P.F., & Henry, N.W. (1968). *Latent structure analysis.* Boston, MA: Houghton Mifflin.

Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology, 95,* 211–224. doi:10.1037/0022-0663.95.2.211

Leslie, L., & Caldwell, J. (2011). Qualitative Reading Inventory-V (QRI-V). *Boston, MA: Pearson.*

Lyon, G. R., & Fletcher, J. M. (2001). Early identification, prevention and early intervention for children at-risk for reading failure. *Publication of the Council for Basic Education*.

Masyn, K. (2013). Latent class analysis and finite mixture modeling. In T. Little (Ed.), *The Oxford handbook of quantitative methods in psychology* (Vol. 2, pp. 375–393). Oxford, UK: Oxford University Press.

Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Reading Research Quarterly*, *40*(2), 148-182. doi: 10.1598/rrq.40.2.2

McLachlan, G., & Peel, D. (2004). *Finite mixture models*. John Wiley & Sons.

Meisinger, E. B., Bradley, B. A., Schwanenflugel, P. J., Kuhn, M. R., & Morris, R. D. (2009). Myth and reality of the word caller: The relation between teacher nominations and prevalence among elementary school children. *School Psychology Quarterly*, 24, 147.

Muthén, B. (2003). Statistical and substantive checking in growth mixture modeling: comment on Bauer and Curran (2003). *Psychological Methods, 8,* 369-377. doi: 10.1037/1082-989X.8.3.369

Nation, K., Adams, J.W., Bowyer-Crane, C.A., & Snowling, M.J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology, 73*, 139-158. doi: 10.1006/jecp.1999.2498

Nation, K. Cocksey, J., Taylor, J.S.H., & Bishop, D.V.M. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *The Journal of Child Psychology and Psychiatry, 51*, 1031-1039. doi: 10.1111/j.1469-7610.2010.0225.x

Nylund, K.L., Asparouhov, T., & Muthén, B.O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14,* 535-569. doi: 10.1080/10705510701575396

Ouellette, G. & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and Writing, 23*, 189-208. doi: 10.1007/s11145-008-9159-1

Pikulski, J. J., & Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. *The Reading Teacher*, *58*(6), 510-519. doi: 10.1598/rt.58.6.2

Rashotte, C.A., MacPhee, K., & Torgesen, J.K. (2001). The effectiveness of a group reading instruction program with poor readers in multiple grades. *Learning Disability Quarterly, 24*, 119-134. doi: 10.2307/1511068

Riedel, B.W. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first-grade students. *Reading Research Quarterly, 42*, 546-567. doi: 10.1598/RRQ.42.4.5

Roth, F.P., Speece, D.L., & Cooper, D.H. (2002). A longitudinal analysis of the connection between oral language and early reading. *The Journal of Educational Research, 95*, 259-

272. doi: 10.1080/00220670209596600

Schrank, F. A., Mather, N., & McGrew, K. S. (2014). Woodcock-Johnson IV Tests of Achievement. *Rolling Meadows, IL: Riverside*.

Semel, E. M., Wiig, E. H., & Secord, W. (2006). *CELF 4: clinical evaluation of language Fundamentals*. Pearson: Psychological Corporation.

Silverman, R.D., Speece, D.L., Harring, J.R., & Ritchey, K.D. (2013). Fluency has a role in the simple view of reading. *Scientific Studies of Reading, 17*, 108-133. doi: 10.1080/10888438.2011.618153

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*, 934–947. doi: 10.1037//0012-1649.38.6.934

Torgesen, J.K., Alexander, A.W., Wagner, R. K., Rashotte, C.A., Voeller, K.K.S, Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities, 34*, 33-58. doi: 10.1097/00004703-200108000-00033

Torgesen, J.K., Wagner, R., & Rashotte., C.A. (1999). CTOPP Comprehensive Test of Phonological Processing. *Austin, TX: Pro-Ed.*

Tunmer, W.E., & Chapman, J.W. (2012). The simple view of reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities,45*, 453-466. doi: 10.1177/0022219411432685

Vellutino, F.R., Tunmer, W.E., Jaccard, J.J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading, 11*, 3-32. doi: 10.1080/10888430709336632

Vermunt, J. K. (2010). Latent Class Modeling with Covariates: Two improved Three-Step Approaches. *Political Analysis*, *18*, 450–469. doi:10.1093/pan/mpq025

Wagner, R.K. & Torgesen, J.K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101,* 192-212.

Wagner, R.K., Torgesen, J.K., Rashotte, C.A., Hecht, S.A., Barker, T.A., Burgess, S.R., Donahue, J., Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5 year longitudinal study. *Developmental Psychology*, *33*, 468-479. doi: 10.1037//0012-1649.33.3.468

Wiederholt, J. L., & Bryant, B. R. (2012). Gray Oral Reading Test–Fifth Edition (GORT-5): Examiner's manual. *Austin*, *TX: Pro-Ed.*

Table 1

Descriptive Statistics Disaggregated by Risk Status and t-tests

	At-risk					Not At-risk					
Measure	*n*	*M*	*SD*	Min	Max	*n*	*M*	*SD*	Min	Max	*t*
Elision	141	6.82	2.66	1.00	16.00	147	10.35	2.68	3.00	19.00	11.20
Blending Words	141	9.34	2.07	4.00	13.00	147	11.53	2.47	4.00	18.00	8.15
Letter Word	141	77.19	12.17	40.00	112.00	148	100.34	14.41	70.00	135.00	14.78
Word Attack	141	83.82	16.88	42.00	124.00	148	107.40	11.60	73.00	132.00	13.90
CELF	141	6.58	3.74	1.00	14.00	147	9.12	3.43	1.00	15.00	6.00
QRI	141	2.44	1.59	0.00	6.00	149	3.57	1.67	0.00	6.00	5.90
GORT rate	140	5.39	1.21	2.00	11.00	148	9.01	2.60	4.00	15.00	15.30
WJ PC	140	81.76	10.96	40.00	113.00	148	100.14	11.17	78.00	128.00	14.08
GORT comp	141	4.34	1.02	2.00	11.00	148	8.31	2.53	4.00	14.00	17.32

Note. Scores are norm-referenced except for QRI. CELF = Clinical Evaluation of Language Fundamentals - 4; QRI = Qualitative Reading Inventory - 5; GORT rate = Gray Oral Reading Tests - 5 Rate subscore; WJ PC = Woodcock Johnson IV Passage Comprehension subtest; GORT comp = Gray Oral Reading Tests - 5 Comprehension subscore; All *t*-tests significant at $p < .001$

Table 2

Fit Statistics of the Latent Profile Analyses by Risk Status

# of classes	LL	# of parameters	BIC	ABIC	AWE	LMR *p*-value
At-Risk						
1	-1204.97	14	2479.23	2434.93	2590.51	-
2	-1103.10	22	2315.07	2245.46	**2489.95**	0.06
3	-1053.88	30	2256.22	2161.30	2494.69	0.21
4	-1009.64	38	2207.33	2087.10	2509.39	0.26
5	Non-positive definite					
Not At-Risk						
1	-1367.71	14	2805.47	2761.17	2917.53	-
2	-1223.55	22	2557.18	2487.56	**2733.27**	0
3	-1172.44	30	2495.00	2400.06	2735.12	**0.01**
4	-1141.56	38	**2473.27**	2353.01	2777.42	0.22
5	-1122.98	46	2476.15	2330.57	2844.32	0.43

Note. Bold values indicate the preferred model for a given fit statistic. LL = Log-likelihood; BIC = Bayesian Information Criterion; ABIC = Adjusted BIC; AWE = Approximate Weight of Evidence; LMR = Lo-Mendell-Rubin Likelihood Ratio Test.

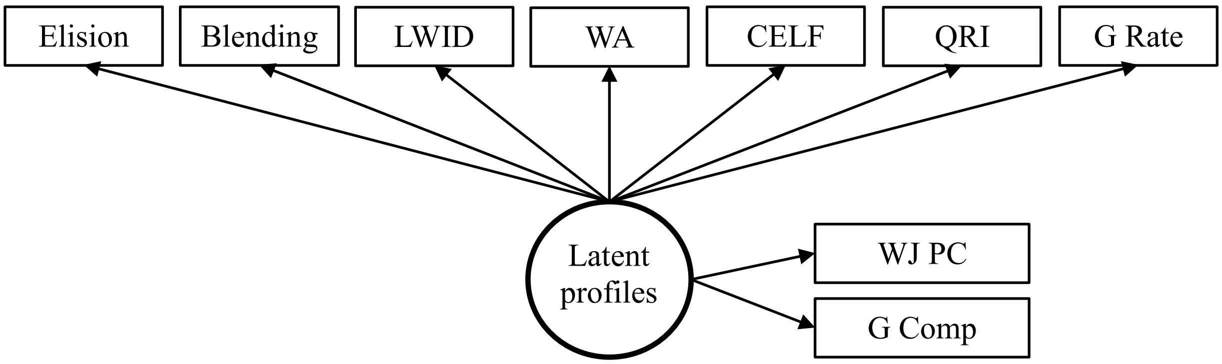

Figure 1. Conceptual diagram of latent profile analyses for both the at-risk and not at-risk groups. LWID = Letter Word Identification; WA = Word Attack; CELF = Clinical Evaluation of Language Fundamentals – 4 Understanding Paragraphs subtest; QRI = Qualitative Reading Inventory - 5; G rate = Gray Oral Reading Tests - 5 Rate subscore; WJ PC = Woodcock Johnson IV Passage Comprehension subtest; G comp = Gray Oral Reading Tests - 5 Comprehension subscore.

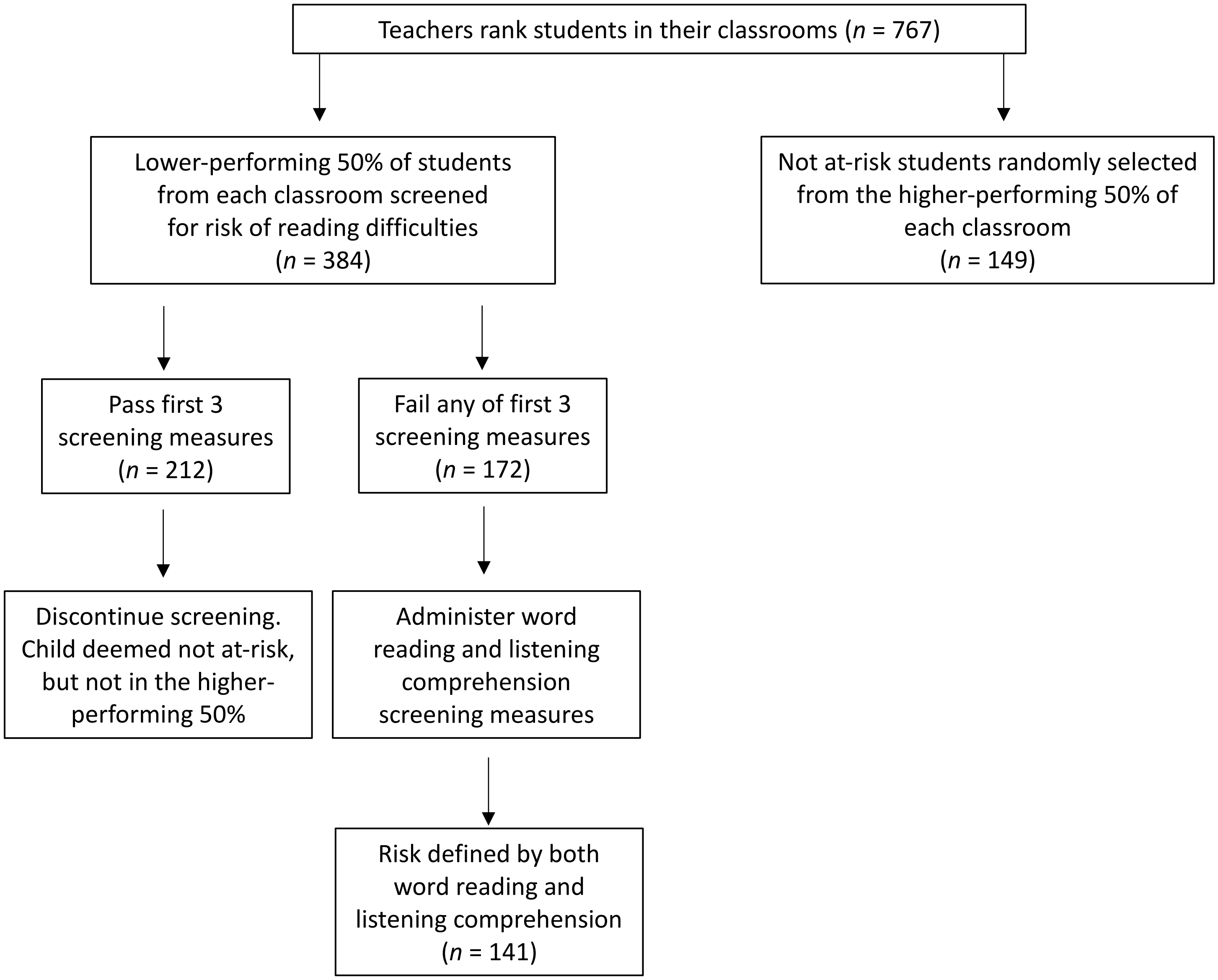

Figure 2. Flow chart summarizing the process used to determine at-risk status and choosing the not at-risk subsample.

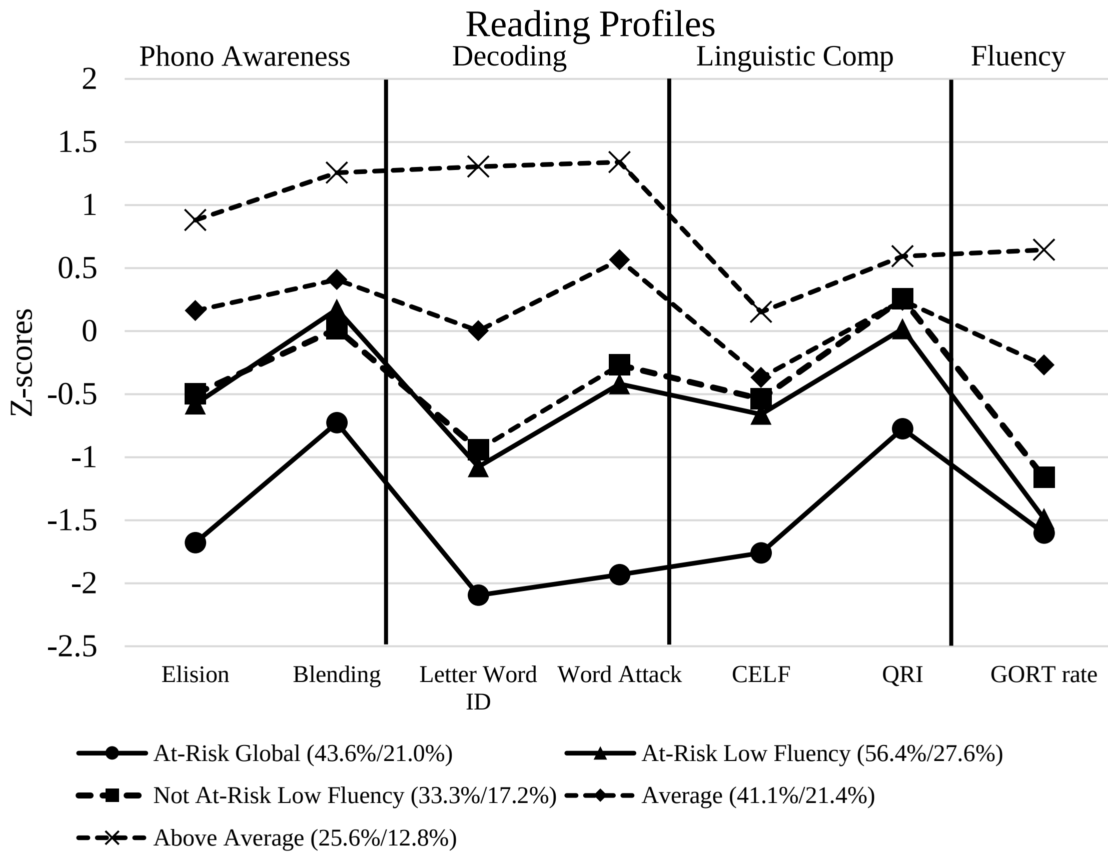

Figure 3. Item-profile plot of at-risk and not at-risk students. Separate LPAs were conducted for each subsample, but are presented on one plot for ease of interpretation. Solid lines denote latent profiles of at-risk students. Dashed lines denote latent profiles of not at-risk students. Percentages before the slash represent the percent within each risk status group. Percentages after the slash represent the percent of the overall sample.

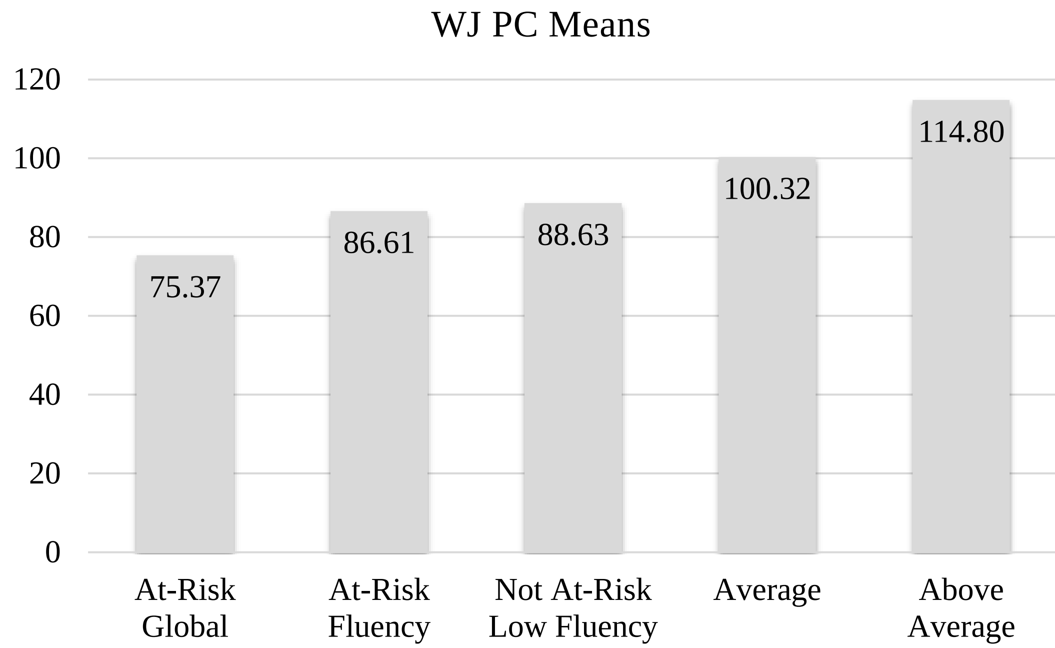

Figure 4. Profile-specific means of WJ-IV Passage Comprehension. All pairwise comparisons were significantly different except for the means of the At-Risk Fluency and Not At-Risk Fluency profiles (p = .25).

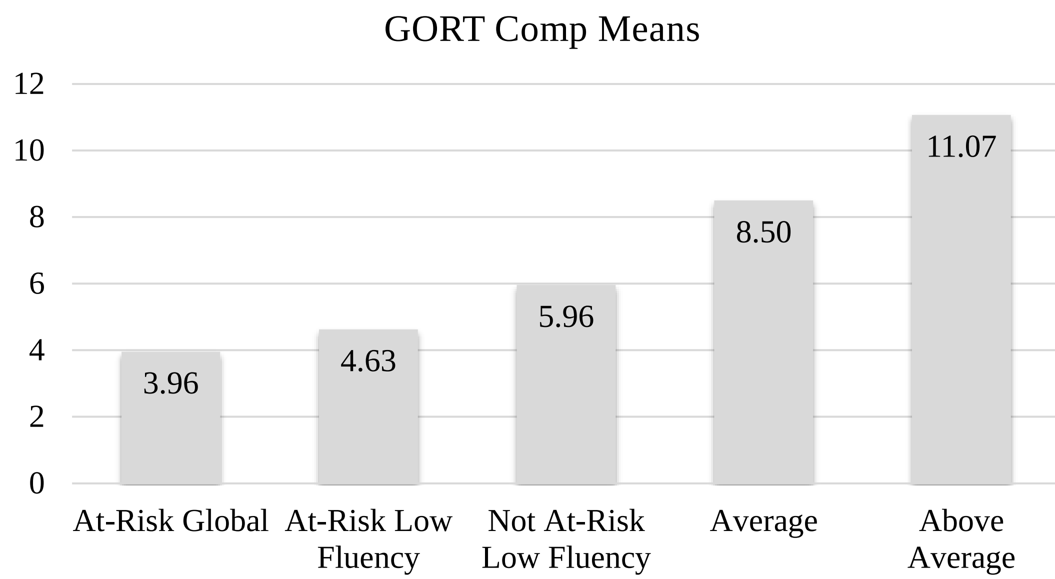

Figure 5. Profile-specific means of GORT Comprehension. All pairwise comparisons were significantly different.