Published in final edited form as:
*J Learn Disabil*. 2012 ; 45(3): 217–231. doi:10.1177/0022219412442154.

# First-Grade Cognitive Abilities as Long-Term Predictors of Reading Comprehension and Disability Status

**Douglas Fuchs, PhD**[1], **Donald L. Compton, PhD**[1], **Lynn S. Fuchs, PhD**[1], **V. Joan Bryant, PhD**[1], **Carol L. Hamlett, MS**[1], and **Warren Lambert, PhD**[1]

[1]Vanderbilt University, Nashville, TN, USA

## Abstract

In a sample of 195 first graders selected for poor reading performance, the authors explored four cognitive predictors of later reading comprehension and reading disability (RD) status. In fall of first grade, the authors measured the children's phonological processing, rapid automatized naming (RAN), oral language comprehension, and nonverbal reasoning. Throughout first grade, they also modeled the students' reading progress by means of weekly Word Identification Fluency (WIF) tests to derive December and May intercepts. The authors assessed their reading comprehension in the spring of Grades 1–5. With the four cognitive variables and the WIF December intercept as predictors, 50.3% of the variance in fifth-grade reading comprehension was explained: 52.1% of this 50.3% was unique to the cognitive variables, 13.1% to the WIF December intercept, and 34.8% was shared. All five predictors were statistically significant. The same four cognitive variables with the May (rather than December) WIF intercept produced a model that explained 62.1% of the variance. Of this amount, the cognitive variables and May WIF intercept accounted for 34.5% and 27.7%, respectively; they shared 37.8%. All predictors in this model were statistically significant except RAN. Logistic regression analyses indicated that the accuracy with which the cognitive variables predicted end-of-fifth-grade RD status was 73.9%. The May WIF intercept contributed reliably to this prediction; the December WIF intercept did not. Results are discussed in terms of a role for cognitive abilities in identifying, classifying, and instructing students with severe reading problems.

### Keywords

cognitive abilities; reading disability identification; reading comprehension

## First-Grade Cognitive Abilities as Long-Term Predictors of Reading Comprehension and Disability Status

In 2009, two thirds of fourth graders across America were reading below Proficient (http:nationsreportcard.gov/reading_2009), the performance goal set by the National Assessment Governing Board. If this weren't troubling enough, we know from recent research on responsiveness-to-instruction (RTI) that many students' poor reading performance is unaffected by the best and most intensive instruction researchers can deliver in their field-based studies and university clinics. Researchers estimate that these low



**Corresponding Author:** Douglas Fuchs, Vanderbilt University, 228 Peabody, Nashville, TN 37203, USA doug.fuchs@vanderbilt.edu.

**Declaration of Conflicting Interests**
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

responders—including children with formally diagnosed reading disabilities (RD)—represent as much as 5% of the school-age (K–12) population (e.g., Torgesen, 2000; Wanzek & Vaughn, 2009), or about 2.5 million children and youth. Moreover, when practitioners rather than researchers try to implement best-evidence reading practices, we estimate that this proportion of low responders doubles to 10% of the school-age population, or 5 million students. Although this estimate is more an educated guess than fact, we are confident nonetheless that there will be many more such children when practitioners are the instructors because they are less likely than researchers to implement the instruction with fidelity, which we believe is often a function of insufficient training and time for instruction and related resource issues.

The 5% and 10% problems can help us think about the large number of students reading below proficiency on the National Assessment of Educational Progess (NAEP). The first percentage signals that the research community, for all its hard work and achievements, does not yet know how to teach reading to all students in this country. The second percentage suggests the difficulty inherent in bridging research and practice (e.g., D. Fuchs, Fuchs, Harris, & Roberts, 1996). Our lack of knowledge about how to teach reading to the low responders and our lack of knowledge about how to scale-up effective instructional programs are important, complex, and separate issues. Solving one doesn't solve the other.

In this article, we explore the 5% problem by asking how researchers and practitioners may identify and help the children for whom best-evidence practices in reading are not enough. Toward this end, we use a longitudinal database to look at whether and how well the cognitive abilities and reading development of poor readers in first grade predict their performance on a measure of reading comprehension and their RD status at fifth grade. A basic belief and motivation for this study is that as the research community becomes more knowledgeable about the cognitive abilities of these most at-risk students, it will understand how to identify them more accurately and efficiently and teach them more effectively. Before describing our methods and results, we provide background for our study purposes.

## Reading and Cognition

**Beginning reading—**More than three decades of research has established that young children must acquire decoding and word recognition skills to become independent readers (e.g., Chall, Jacobs, & Baldwin, 1990; Ehri, 1998; Rieben & Perfetti, 1991). Researchers have also identified cognitive characteristics responsible for, or strongly associated with, children's failure to read at the word level. RD is understood by many as almost always the result of deficits in phonological awareness, speeded lexical retrieval, and verbal short-term memory (e.g., Rack, Snowling, & Olson, 1992; Vellutino, Fletcher, Snowling, & Scanlon, 2004). The most important of these processes is phonological awareness (e.g., Castles & Coltheart, 2004; Hulme, Snowling, Caravolas, & Carroll, 2005). Children with RD use phonological awareness to read, but they are much less skillful in their use of it than are their typically developing peers (e.g., Rack et al., 1992; Ramus & Szenkovits, 2009).

Speeded lexical retrieval or naming speed deficits rarely occur without phonological difficulty (Compton, DeFries, & Olson, 2001; Schatschneider, Carlson, Francis, Foorman, & Fletcher, 2002; Vukovic & Siegel, 2006), but they have separate and unique predictive power (Catts, Gillespie, Leonard, Kail, & Miller, 2002; Manis, Doi, & Bhadha, 2000; Wolf & Bowers, 1999). So does verbal short-term memory, which is understood to be a component of the working memory system. Researchers have found that performance on tasks tapping retrieval in verbal short-term memory (de Jong, 1998; Fletcher, 1985; Ramus & Szenkovits, 2009) contribute unique variance to the prediction of word reading skill, even when controlling for phonological awareness or naming speed (Ramus & Szenkovits, 2009; Swanson & Howell, 2001).

**Later reading—**Most researchers agree that the importance of word reading and language skills shifts over time, with word recognition contributing more variance in earlier development and language explaining more variance later (e.g., Catts, Hogan, & Adlof, 2005; Francis, Fletcher, Catts, & Tomblin, 2005; Gough, Hoover, & Peterson, 1996; Hulslander, Olson, Willcutt, & Wadsworth, 2010). However, in contrast to what is known about skill development and RD in the primary grades, there seems to be less certainty about both in the intermediate grades. Although some have reported findings that RD is the result of a specific comprehension deficit, rather than word reading deficit (e.g., Badian, 1999), others have described more complex profiles of such students. As described by Compton, Fuchs, Fuchs, Elleman, and Gilbert (2008), Leach, Scarborough, and Rescorla (2003) studied 31 students with RD in Grades 4 and 5. They described 11 as poor decoders, 10 with a specific comprehension deficit, and 10 with both decoding and comprehension problems. Thus, more than two thirds of Leach et al.'s sample had at least some difficulty with word identification and decoding, a finding corroborated by Lipka, Lesaux, and Siegel (2006).

Moreover, relatively little is known about the cognitive abilities associated with comprehension skills, especially in comparison to extant knowledge about word reading skills. Reading competence in the intermediate grades largely involves comprehending text, which may depend on oral language abilities, including vocabulary knowledge (Dickinson et al., 2003; Muter, Hulme, Snowling, & Stevenson, 2004; Oakhill, Cain, & Bryant, 2003; Scarborough, 2005; Sénéchal, Ouellette, & Rodney, 2006) and semantic knowledge (Catts, Fey, Zhang, & Tomblin, 2001; Leach et al., 2003; Nation, Marshall, & Snowling, 2001; Nation & Snowling, 1998, 1999). Another important cognitive factor may be nonverbal problem solving, which requires students to analyze relations among and draw inferences about characters or actions in narrative text and to decipher challenging expository material. The contribution of nonverbal problem solving to reading comprehension in the intermediate grades, however, has been infrequently explored.

Even less is known about the early cognitive determinants of reading comprehension and RD in the intermediate grades, reflecting the infrequency with which longitudinal studies have been conducted. Exploring cognition in this context is potentially important because findings may facilitate development of instructional programs that mitigate or preclude later comprehension difficulties and because they may contribute to the production of assessments that help identify children with RD sooner than later.

## A Two-Stage Assessment Process to Identify Nonresponders

Beyond exploring the importance of first-grade cognitive abilities and reading performance to reading comprehension and RD status in fifth grade, we conducted this study to contribute to the development of effective and efficient RTI frameworks. According to many (e.g., American Psychological Association, 2005; Council for Exceptional Children, 2007; National Joint Committee on Learning Disabilities, 2005), RTI has two important goals, both of which rely on the use of valid assessment. The first goal encourages educators to identify at-risk students for early intervention and prevention, the second to recognize those for whom such early intervention is ineffective and to ensure that they get more intensive instruction and possibly special education.

In most RTI frameworks, practitioners rely on a universal screen to identify children for early intervention. Unfortunately, research indicates most screening instruments overidentify at-risk students (e.g., D. Fuchs, Compton, Fuchs, & Davis, 2008; Seethaler & Fuchs, 2010). For this reason, we (e.g., Compton et al., 2010; D. Fuchs, Compton, Fuchs, Bouton, & Caffrey, 2011; L. Fuchs et al., in press) and others have recommended a two-stage assessment process, whereby schools use screening instruments to exclude "true negatives" (i.e., students clearly *not* at risk) and involve the remaining students in a second-stage assessment to more accurately distinguish true positives from false positives. Such a process can help schools restrict the use of costly interventions to only those students in need of them and, in this way, contribute to efficient RTI frameworks. In this investigation, we explored the utility of a two-stage process in which a universal screening instrument was our Stage 1 assessment and measures of cognitive abilities and reading progress was our Stage 2 assessment.

RTI's second goal, as mentioned, is to identify the subset of chronically low responders who require more intensive and sustained intervention, similar to the individualized programming often recommended for students with RD (e.g., D. Fuchs & Fuchs, 2006). Some RTI frameworks operationalize RD among young children in terms of their inadequate response to the first-grade classroom instruction. In the present study, end-of-first-grade reading performance might in like manner be considered an appropriate index of RD classification. At the same time, it is interesting to consider whether these chronic low responders might be identified earlier, in fall of first grade, partly on the basis of the cognitive abilities or December reading performance, as suggested elsewhere in this issue (Compton, Gilbert, et al., in press). Earlier identification of such children, again, potentially contributes to more effective and efficient RTI frameworks.

### Purpose

In the present study, we attempted to predict reading comprehension in spring of fifth grade for students on whom we had rapid automatized naming, phonological processing, oral language comprehension, and nonverbal reasoning data from fall of their first grade year. In addition, we modeled the children's first-grade reading development to derive December and May intercepts to explore the predictive value of the cognitive variables when in competition, so to speak, with first-grade reading. Put differently, by including first-grade reading development as a predictor, we attempted to control for cognitive abilities involved in word-level skill as we considered the predictors of fifth-grade reading comprehension. We also contrasted the amount of variance explained in fifth-grade reading comprehension by the cognitive versus reading performance variables as a function of when first-grade reading skill was estimated. Finally, we used logistic regression to explore the accuracy with which the cognitive variables predicted end-of-fifth-grade RD status with and without the reading performance data.

## Method

### Participants

Participants were selected from 42 first-grade classrooms in 16 schools in two school districts—one urban and the other suburban—in middle Tennessee. Eight of the schools were supported by Title I dollars. From the 42 classes, the 6 poorest readers in each class were selected, totaling 252 low-performing children. To select these students, project staff administered Word Identification Fluency (WIF; L. Fuchs, Fuchs, & Compton, 2004) and Rapid Letter Naming (RLN) from the *Comprehensive Test of Phonological Processing* (CTOPP; Wagner, Torgesen, & Rashotte, 1999) to all "consented" children, low performing or otherwise. A total of 783 students were screened, representing 90% of all children in the 42 classrooms. For descriptions of WIF and RLN, see the Measures section.

In each class, children were rank ordered on the two tests, and the six lowest performing children were selected for study inclusion. Project staff presented teachers with the names of these students as well as the names of three alternate children. More than 95% of the time, teachers concurred with project staff's selection of the six lowest readers. When there was disagreement, the child in question was replaced by the teacher, who chose among the three alternates. Participant selection occurred in late September and early October. This sample of low-performing children was followed longitudinally from fall of first grade through spring of fifth grade. During the 5 years, we tracked and tested annually 195 of the initial sample of 252, reflecting an attrition rate of 22.6%, or an average 4.5% per year. The 195 “stayers” and 57 “movers” were comparable on the WIF and RLN measures. See Table 1 for demographic and performance data on the full sample of 195 “stayers” (as well as on the non-RD and RD subsamples, described later).

Among the “stayers,” 131 (67.2%) participated in small-group tutoring for 10 weeks in fall or spring of first grade (D. Fuchs et al., 2008). Its purpose was to strengthen the children’s word recognition skills. Because there were no long-term effects of this tutoring, and because the disability status of the 131 children in spring of fifth grade (see below) was not affected by whether they were tutored, we do not describe our tutoring program here.

## Measures

We explain measures used (a) to screen children for study entry, (b) to assess their first-grade cognitive characteristics with which we predicted their long-term reading performance and disability status, and (c) to model their reading development during first grade and their reading comprehension in spring of fifth grade.

**Universal screening to identify low-performing children—**As indicated, we used WIF and RLN to identify the initial 252 first-grade study participants. With WIF (L. Fuchs et al., 2004), children are presented with a single page of 50 high-frequency words randomly sampled from 100 high-frequency words from the Dolch preprimer, primer, and first-grade-level lists. They have 1 min to read as many words as they can. If they hesitate on an item for 4 s, the examiner prompts them to proceed. If they finish reading in less than 1 min, the score is prorated. As part of our screening, the children were directed to read not one but two different 50-word lists and their WIF score was an average of their performance on each. Test–retest reliability was .85 for all those we screened.

RLN (Wagner et al., 1999) requires children to name six letters arranged in random order on two pages. Each page displays four rows with nine letters per row. The tester asks the child to name each letter, corrects any errors, and then asks the child to name the letters again. If the child cannot name all of them or makes more than four errors during testing, the examiner discontinues testing. The child’s score is the number of seconds required to name the 36 letters. Test–retest reliability is .97 for 5- to 7-year-olds.

**Assessing first-grade cognitive characteristics—**We explored the predictive utility of four first-grade cognitive characteristics or dimensions: Rapid Automatized Naming (RAN), phonological processing, oral language comprehension, and nonverbal reasoning. For each of these, we administered multiple measures and created factor scores to strengthen their respective reliabilities. To assess the RAN dimension, we used four CTOPP subtests (Wagner et al., 1999), the first of which was RLN, described above. The format and procedures for the other three RAN measures parallel that of RLN. For Rapid Digit Naming, the stimulus array includes numerals, and test–retest reliability is .91 for 5- to 7-year-olds. The array for Rapid Color Naming includes colors, and test–retest reliability is .83 for the same age range. For Rapid Object Naming, the array includes pictures of objects, and test–retest reliability is .77 for 5- to 7-year-olds.

We also relied on the CTOPP (Wagner et al., 1999) to describe children’s *phonological processing*, specifically, the subtests of Elision, Sound Matching, Blending Words, Memory for Digits, Non-word Repetition, and Segmenting Words. With Elision, the tester directs the child to say words with a constituent part removed from the words. Test–retest reliability is .

88 for 5- to 7-year-olds. Sound Matching requires children to match first and last sounds in words. For *first* sound matching, children are presented with a word and then are asked which of three words (depicted as pictures) start with the same sound x. A similar procedure explores *last* sound matching. After three practice items, the test comprises 20 items. Test–retest reliability is .83.

With respect to Blending Words, the child listens to a recording of segmented words and is asked to blend them into words. The score is the number of words correctly blended. Test–retest reliability is .88. With Memory for Digits, the child listens to a recording of strings of numbers ranging in length from two to eight digits and repeats each string in the same order. The score is the number of strings repeated without error. Test–retest reliability is .74. With Non-word Repetition, the child listens to a recording of pseudo-words and repeats each one. The score is the number pseudo-words for which all phonemes are produced correctly. Test–retest reliability is .68. With Segmenting Words, the tester says words; the child repeats each word one sound at a time. The score is the number of words for which the child produces all phonemes correctly. Test–retest reliability is .86. With Memory for Words, the tester reads random series of words, which the child repeats in order. The score is the number of series correctly repeated. Coefficient alpha on this sample was .84.

Our study participants' *comprehension of oral language* was tested with (a) the Listening Comprehension component of the *Woodcock Diagnostic Reading Battery* (WDRB; Woodcock, 1997), (b) two subtests of the *Wechsler Abbreviated Scale of Intelligence* (WASI; Psychological Corporation, 1999), Vocabulary and Similarities, and (c) the Oral Vocabulary subtest of the *Woodcock–Johnson III Tests of Cognitive Abilities* (WJ III; Woodcock, McGrew, & Mather, 2001). The Listening Comprehension component of the WDRB measures understanding of sentences or passages. Students supply the missing word from the end of each sentence or passage. The test begins with simple verbal analogies and associations and progresses to comprehension involving the ability to discern implications. Testing is discontinued after six consecutive errors. The score is the number of correct responses. Reliability is .80 at ages 5–18; the correlation with the WJ III is .73. Coefficient alpha on this sample was .82.

WASI Vocabulary measures expressive vocabulary, verbal knowledge, and foundation of information. The first four items present pictures. The child identifies the object in each picture. For remaining items, the tester says a word, which the child is expected to define. Responses are awarded 0, 1, or 2 points, depending on quality of response. Testing is discontinued after five consecutive scores of 0. The score on this measure is the total number of points earned. Splithalf reliability is .86 to .87 at ages 6–7 (Zhu, 1999). Coefficient alpha on this sample was .83. WASI Similarities measures verbal concept formation, abstract verbal reasoning ability, and general intellectual ability. For the first four items, the tester presents two rows of objects. The child finds the object in the bottom row that is most similar to items in the top row. For the remaining items, the tester says two words and the child identifies how the words are alike. Testing is discontinued after four consecutive errors. The score is the number of correct items. Split-half reliability is .88 to . 89 at ages 6–7 (Zhu, 1999). Coefficient alpha on this sample was .88. WJ III Oral Vocabulary assesses the ability to provide synonyms and antonyms in response to stimulus words presented by the examiner. Split-half reliability on the present sample was .92.

To assess *nonverbal reasoning*, we used WASI Block Design and WASI Matrix Reasoning. Block Design measures spatial visualization, visual-motor coordination, and abstract conceptualization abilities. The tester presents three-dimensional block models or two-dimensional printed models; the child replicates the design with blocks. Testing is discontinued after three consecutive errors. The score is the number of correct replications.

Split-half reliability is .84 to .85 at ages 6–7 (Zhu, 1999). Coefficient alpha on this sample was .83. Matrix Reasoning measures nonverbal fluid reasoning. The tester presents a series of patterns and the child selects “missing pieces” from five choices. Testing is discontinued after four errors on five consecutive items. The score is the number of correct responses. Split-half reliability is .94 to .96 at ages 6–7 (Zhu, 1999). Coefficient alpha on this sample was .91.

**Modeling first-grade reading progress and fifth-grade reading outcomes—**To model reading development during first grade, we administered two alternate forms of WIF each week for 18 weeks, beginning in November and ending in April. As already indicated, the two weekly scores were averaged. To model a reading outcome in spring of fifth grade, we administered the Passage Comprehension subtest of the *Woodcock Reading Mastery Tests–Revised* (WRMT-R; Woodcock, 1998) each spring in Grades 1 through 5. The Passage Comprehension subtest uses a modified cloze (or maze) procedure. For the first set of items, the tester presents a rebus, and the child points to the picture corresponding to it. Next, the child points to the picture representing words printed on the page. On later items, the child reads a passage silently and identifies the missing word. Split-half reliability is .90.

### Procedure

In September of first grade, we screened students for study participation. In October, we administered the cognitive measures to each child individually in three sessions, requiring about 110 min across the sessions. Beginning in November, students’ reading development was assessed weekly. Across the school year, we collected 18 weeks of WIF data—9 weeks in both fall and spring. In spring of Grades 1–5, we administered the Passage Comprehension subtest of the WRMT-R. Prior to each testing wave, staff learned, practiced, and established agreement on test administration procedures. If a tester failed to achieve a criterion of 90% accuracy in administering each test to the project coordinator, who acted as a student, the tester received additional training on the relevant measure and then completed another accuracy assessment until agreement exceeded 90%. All test sessions were audiotaped, and a second scorer checked all tapes to identify scoring errors, which were corrected.

## Data Analysis and Results

### Preliminary Analyses

**First-grade cognitive predictors: Missing data and data reduction—**Only 20 of 3,120 observations associated with the cognitive variables were missing (< 1%). We eliminated the missing values by using single imputation with an “expectation maximization” algorithm (Rubin, 1991). The algorithm preserves the sample’s variance in contrast to mean or median imputation methods, which unfairly lower it. The standard deviations of the imputed cognitive predictor variables in this sample averaged 101% of the standard deviation before imputation.

We grouped children’s responses on our various cognitive measures into conceptually related cognitive dimensions (i.e., RAN, phonological processing, comprehension of oral language, nonverbal reasoning) and conducted confirmatory factor analysis with PROC CALIS (Hatcher, 1994) in SAS 9.2 (SAS Institute, 2008). The model posited a simple structure with each cognitive dimension measuring only one factor. To evaluate it, we used one residual index and one fit index (see Hu & Bentler, 1999), and we relied on Hu’s fit criteria (see Tabachnick & Fidell, 2007), that is, Bentler’s (1990) comparative fit index (Hu’s cutoff value of about .95) and the standardized root mean square residual (Hu’s cutoff value of about .08). The fit indices for our sample’s confirmatory factor analysis were comparative fit index = .94 and standardized root mean square residual = .06. Confirmatory factor analysis also revealed that standardized loadings on the scores ranged from .45 to .86.

We then calculated cognitive dimension scores as the *z* average of each factor's items to assist others' replication (Dawes, 1979). These cognitive dimension scores correlated well with actual factor scores from PROC CALIS (RAN $r = 1.00$; phonological processing $r = .95$; oral language comprehension $r = .97$; nonverbal reasoning $r = .96$), suggesting the *z*-averaged cognitive dimension scores are useful approximations of proper, model-based factor scores. Thus, we treated the four cognitive dimension scores as useful simplifications of the four factor scores. Correlations among the cognitive dimension scores ranged from −.15 to .54 (see the first three columns of Table 2), suggesting that they were correlated but distinct from each other. As expected, the RAN factor (i.e., time to name items) was negatively correlated with the other three factors.

**Modeling first-grade reading progress and fifth-grade reading outcome—**We had 18 weekly WIF scores in first grade, as well as year-end Passage Comprehension measurements in Grades 1 through 5. For WIF, we were interested in modeling a December outcome and a May outcome; for Passage Comprehension, we wanted a fifth-grade outcome. We defined outcome as the final intercept of a child's individual linear growth curve (Rogosa, Brandt, & Zimowski, 1982).

To estimate final intercept, we analyzed individual growth curves with a longitudinal mixed model (Singer & Willett, 2003; i.e., hierarchical linear model—see Raudenbush & Bryk, 2002). We ran two models on WIF raw scores to derive final intercepts in December (Week 6 of WIF data collection) and May (Week 18). We also ran a model on Passage Comprehension raw scores to obtain a fifth-grade final intercept. In each of these models, the intercept was set to represent performance at the end point, estimating a best fit linear individual growth curve to the child's available scores.

A likelihood-ratio test was applied to determine whether adding slope resulted in a better fit than a means model alone (Singer & Willett, 2003). For December WIF, the difference in –2 log likelihood for means versus means + slope was significant, $\chi^2(1, N = 992) = 339.50$, $p < .0001$, suggesting that the inclusion of slope made a better fit for the model. The same was true both for May WIF, $\chi^2(1, N = 3274) = 285.60$, $p < .0001$, and fifth-grade Passage Comprehension, $\chi^2(1, N = 884) = 531.30$, $p < .0001$. For each of the three models, we assessed the reliability of the intercept, which in HLM is calculated by dividing the true variance by the true variance plus error variance. For the December WIF intercept, May WIF intercept, and fifth-grade Passage Comprehension intercept, the respective reliabilities were .90, .86, and .84.

**Designating RD status—**For fifth-grade Passage Comprehension, we also set cut points to designate RD and non-RD (NRD) status. We expressed each student's final intercept as a standard score, which was derived from normative scores on the Passage Comprehension subtest of the WRMT-R (Woodcock, 1998; $M = 100$, $SD = 15$). Students scoring 85 and below were designated RD; 92 and above, NRD. Students who scored below 92 and above 85 were in a "buffer zone" and were eliminated from analyses.

Of the 195 low-performing students in our sample, 36 were RD in spring of fifth grade. This was 4.6% of the 783 students who were screened for study entry in fall of first grade. Among these 36 students with RD, 16.7% had an individualized education program (IEP) in first grade; 39.1% had an IEP with a reading goal in spring of fifth grade. By contrast, 5.1% of 98 NRD students had an IEP in first grade; by spring of fifth grade, only 1 (1.0%) had an

IEP with a reading goal. See Table 1 for demographic and performance data by children's RD and NRD status.

## Predicting Fifth-Grade Reading Outcome

See Table 2 for correlations among the four first-grade cognitive predictors and first-grade December and May WIF intercepts and the fifth-grade Passage Comprehension intercept (or outcome). We used four models to predict the fifth-grade Passage Comprehension final intercept. Each incorporated the four cognitive predictors. Two of the four models also included the December WIF intercept as a predictor; the remaining two models used the May, rather than December, intercept as a predictor.

In conducting regression analysis with December WIF scores and the four cognitive dimensions, we first entered the December WIF intercept, $R^2 = .250$, SEE = 3.14, $F$ Change(1, 193) = 65.66, $p < .001$, and next added the four cognitive dimensions, $R^2 = .503$, SEE = 2.56, $R^2$Change = .262, $F$ Change(4, 189) = 25.51, $p < .001$. We then reversed this order, with the cognitive predictors entered first, $R^2 = .449$, SEE = 2.72, $F$ Change(4, 190) = 38.70, $p < .001$, and the WIF intercept added next, $R^2 = .503$, SEE = 2.56, $R^2$Change = .066, $F$ Change(1, 189) = 25.94, $p < .001$. Together, these cognitive and reading predictors accounted for 50.3% of the variance in the fifth-grade Passage Comprehension intercept, $F(5, 189) = 40.21$, $p < .001$. Partitioning this variance revealed that 34.8% of the explained variance was shared between the December WIF intercept and the cognitive dimensions, 13.1% was unique to the December WIF intercept, and 52.1% was unique to the cognitive predictors. In Table 3, we show *B, SE*, β, *t*, and *p* values for the constant and each predictor. All five predictors contributed unique variance to the prediction of the fifth-grade Passage Comprehension outcome.

We conducted a parallel set of analyses incorporating the May WIF intercept as a predictor. We first entered the May WIF intercept, $R^2 = .407$, SEE = 2.80, $F$ Change(1, 193) = 132.47, $p < .001$, and then added the cognitive dimensions, $R^2 = .621$, SEE = 2.26, $R^2$Change = .214, $F$ Change(4, 189) = 26.59, $p < .001$. We then reversed the order, with the cognitive dimensions entered first, $R^2 = .437$, SEE = 2.72, $F$ Change(4, 190) = 38.71, $p < .001$, and the May WIF intercept added next, $R^2 = .621$, SEE = 2.26, $R^2$Change = .172, $F$ Change(1, 189) = 85.45, $p < .001$. Together, these cognitive and reading predictors accounted for 62.1% of the variance in the fifth-grade Passage Comprehension intercept, $F(5, 189) = 61.81$, $p < .001$. Partitioning this variance revealed that 37.8% of the explained variance was shared between the May WIF intercept and the cognitive dimensions, 27.7% was unique to the May WIF intercept, and 34.5% was unique to the cognitive predictors. As shown in Table 3, all cognitive predictors, except RAN, contributed unique variance to the prediction of the Grade 5 Passage Comprehension outcome.

## Predicting Fifth-Grade RD Status

We used logistic regression to predict RD or NRD status in spring of Grade 5. Despite that the just-described analyses of continuous data should have a higher $R^2$, we conducted logistic regression because it has practical application to RTI. Risk status is often determined in fall of first grade to facilitate early intervention for children likely to develop problems in reading. We explored the utility of five logistic regression models. In Model 1, we relied exclusively on the December WIF intercept. In Model 2, we combined the December WIF score with the four cognitive predictors. We relied on only the May WIF intercept in Model 3. In Model 4, we combined the May WIF score with the four cognitive predictors. In Model 5, we included only the four cognitive predictors.

As in RTI, we were interested in maximizing true positives (i.e., first graders identified as at risk and who truly required reading intervention) and limiting false positives (i.e., first graders identified as at risk but who completed Grade 5 above the RD cut score). Therefore, in all five logistic regression models, we held sensitivity at 91.7. This signified that no more than 3 of the 36 students with RD in spring of fifth grade were identified as NRD at the time of risk specification. Then, we observed how the competing models affected specificity. We used measures of sensitivity, specificity, overall hit rate, and area under the receiver operating curve (ROC) to contrast the utility of the five models (see Note 1).

To contrast the predictive accuracy of the logistic regression models, we used area under the ROC curve (AUC) as a measure of discrimination (see Swets, 1992). To illustrate its use, if we had already placed children into their correct RD or NRD groups and then selected one child at random from each group, we would assume that the child scoring higher on the screening measure(s) would be the NRD child. The AUC represents the proportion of randomly chosen pairs of students for which the screener(s) correctly classified children with and without RD. The greater the AUC, the less likely that classification is the result of chance. It ranges from .50 to 1.00. AUC below .70 indicates a poor predictive model; .70 to .80, fair; .80 to .90, good; and greater than .90, excellent (Swets, 1992).

In Table 4, we report results of these logistic regression analyses. The AUC for the five models ranged from .791 to .910, which is fair to excellent. Models 1, 2, and 5 included only the fall predictors; Model 3, only the spring predictors; and Model 4, the fall and spring predictors. Applying Model 1 (i.e., December WIF intercept as the sole predictor) resulted in 26.5% specificity when sensitivity was held at 91.7%. The December WIF intercept was statistically significant in the model. Adding the four cognitive variables to the December WIF score (Model 2) increased specificity to 67.3%. RAN, phonological processing, and nonverbal reasoning all contributed significantly, and model fit was superior to that of Model 1, $\chi^2(4, N = 134) = 29.67$, $p < .0001$. The four cognitive predictors without a WIF intercept (Model 5) produced identical specificity (67.3%), with the same three cognitive dimensions contributing significantly. Because model fit for the December WIF intercept plus the cognitive variables (Model 2) was not significantly different from the fit for the model with the cognitive variables alone (Model 5), $\chi^2(1, N = 134) = 2.04$, $p = .153$, Model 5 was seen as more parsimonious and, hence, superior.

Waiting until the end of first grade, with reliance on only the May WIF intercept (Model 3), resulted in 43.9% specificity. The May WIF intercept was significant. When the cognitive predictors were added to the May WIF intercept (Model 4), specificity increased to 68.4%, and the May WIF intercept was the only significant contributor to the prediction. The fit for Model 4 was superior to Model 3 (i.e., the May WIF intercept alone), $\chi^2(4, N = 134) = 20.30$, $p = .0004$, and superior to Model 5, which included only the cognitive variables, $\chi^2(1, N = 134) = 23.22$, $p < .0001$, even though specificity for Models 4 and 5 was not appreciably different (68.4% vs. 67.3%, respectively).

## Discussion

We had two purposes in this study. One was to explore first-grade cognitive predictors of fifth-grade reading comprehension while controlling for fall-of-first-grade, or spring-of-first-grade, reading performance. The second purpose was to determine how well the first-grade

[1] *Sensitivity* (the proportion of children correctly predicted by the model to have RD) is computed by dividing the number of true positives by the sum of true positives and false negatives. *Specificity* (the proportion of children correctly predicted to *not* have RD) is computed by dividing the number of true negatives by the sum of true negatives and false positives. *Overall hit rate* (the proportion of children correctly classified as RD and NRD) represents the overall accuracy of the prediction model. The *area under the ROC curve* (AUC) is a plot of the true positive rate against the false positive rate for different possible cut points of a test.

cognitive abilities and reading performance predicted children with and without RD in spring of Grade 5. Our hope was that by pursuing these purposes we would contribute to the eventual development of better methods of both identification and treatment of children with very serious learning problems.

## Predicting Reading Comprehension and RD Status in Fifth Grade

When using information available in fall of first grade (i.e., four cognitive predictors and December WIF intercept), we accounted for 50.3% of the variance in fifth-grade reading comprehension. Of this explained variance, more was unique to the cognitive variables (52.1%) than the WIF intercept (13.1%). Phonological processing and RAN were statistically significant predictors, suggesting they are important for explaining reading comprehension as well as reading at the word level, as demonstrated by many others.

Oral language comprehension was also a significant predictor, as has been previously demonstrated (e.g., Compton, Fuchs, Fuchs, Lambert, & Hamlett, in press; Dickinson et al., 2003; Leach et al., 2003; Muter et al., 2004; Nation et al., 2001; Oakhill et al., 2003; Scarborough, 2005; Sénéchal et al., 2006). So, too, was nonverbal reasoning. Few studies have examined the role of nonverbal reasoning in reading comprehension despite its apparent importance. Consider, for example, the common-place intermediate-grade task that requires students to analyze relations among and draw inferences about characters or actions in narrative text, or the equally typical task of unpacking expository text loaded with complex, technical meaning. Our findings suggest nonverbal reasoning should be explored by researchers interested in cognitive determinants of reading comprehension in the intermediate grades.

Using the May (rather than December) WIF intercept in the previously-described model, we explained 62.1% (vs. 50.3%) of the variance in fifth-grade reading comprehension, with a beta value of .50 (vs. .30). This was expected because the May intercept was one half year more proximal to the predicted outcome. Similarly, the variance explained by the May WIF intercept was more comparable to—but still less than—the variance explained by the cognitive variables (May WIF intercept, 27.7%; cognitive variables, 34.5%). With the May intercept in our model, RAN was no longer a statistically significant predictor, but phonological processing, oral language comprehension, and nonverbal reasoning continued to be so.

We also predicted disability status because of our interest in a possible role for cognitive characteristics in disability identification. Put differently, a goal in our RTI work has been to develop a two-stage process of at-risk identification that may eventually facilitate more timely and appropriate instruction for high-risk children. In this study, as part of Stage 1, we administered a universal screen (WIF and RLN) in fall of first grade to identify our initial sample of 252 poor readers, which became 195 students with the loss of participants over 5 years. In Stage 2, we administered our cognitive measures in fall of first grade and conducted progress monitoring throughout the school year. To determine the utility of this second-stage assessment, we applied a series of logistic regression (or classification) models after stipulating that no more than three students with RD status in fifth grade would be missed by our first-grade screening procedures.

In our first model, we relied solely on the December WIF intercept. This simple, inexpensive, Stage 2 screen failed to accurately classify many students' RD status. The model's hit rate was 44%; its specificity, 26.5%. In a second model, we added the four cognitive predictors to the December WIF score. This more complex and expensive alternative greatly improved classification accuracy (73.9% hit rate; 67.3% specificity). It also produced a significantly better fit than the model based solely on the December WIF intercept.

Considering that we tested these models on a restricted range of poor readers (rather than on a combination of poor readers and higher-performing readers, or true negatives, who would have boosted classification accuracy), and that these assessment data were available near the start of first grade, our accuracy in predicting RD status 4.5 years later is encouraging. Furthermore, a more parsimonious Model 5, based exclusively on the cognitive predictors, resulted in comparable fit and, hence, may be considered superior to Model 2 (i.e., the December WIF intercept and cognitive predictors).

Delaying prediction to spring of first grade increased the value of our progress-monitoring data, but exclusive reliance on the May WIF intercept, shown in Model 3, resulted in lower specificity compared to that of Model 2 (43.9% vs. 67.3%; see Table 4). Combining this May WIF score with the cognitive predictors (Model 4), however, increased specificity to 68.4%, and the fit for Model 4 was superior to the model based solely on the May WIF intercept. It was also superior to that for Model 5, which comprised the cognitive variables alone. Of course, delaying prediction until the end of first grade is likely also to delay intervention until the following fall.

Across our logistic regression analyses, findings suggest that one can be relatively accurate in predicting RD status at the end of Grade 5 from a cognitive battery administered in fall of first grade—a battery that, as in this study, follows a universal screen of WIF and RLN measures.

## Study Limitations

There are at least five important study limitations and admonitions regarding the generalizability of our results, the first of which is that we recruited a restricted range of poor readers into our study. As shown in Table 1, in fall of first grade our full sample was performing one standard deviation below the mean on most normative tests in our battery. In spring of fifth grade, the full sample's mean standard score on passage comprehension was 91.24 (*SD* = 8.07). The RD students' corresponding score was 79.38 (*SD* = 6.61). Even students classified as NRD achieved a mean raw score below that of the normative average (*M* = 97.13, *SD* = 4.71). Although the selection of our sample is justifiable (i.e., all of the children failed the first-stage screening assessment), researchers recruiting more representative student samples would likely obtain different results.

A second limitation is the arbitrary nature of our RD definition. As described, students with a standard score of 85 and below on the Passage Comprehension subtest of the WRMT-R (Woodcock, 1998) were designated RD; those scoring at or above 92 were NRD. Although many would agree that the RD group's mean standard score was low, at least some would disagree with our cut point and, more generally, with our operationalization of RD. We acknowledge that more or less stringent criteria for RD necessarily affect the utility of one's predictors. Moreover, one may question our choice of the Passage Comprehension subtest of the WRMT-R to model our reading comprehension outcome in spring of fifth grade. The works of Cutting and Scarborough (2006) and Keenan, Betjemann, and Olson (2008) suggest that this measure relies heavily on word reading skills. It may be less aligned with reading comprehension than other measures.

On the other hand, we were able to explore the validity of our reliance on WIF as a measure of first-grade reading in our models. We conducted a post hoc analysis of Model 2 (December WIF intercept plus the four cognitive predictors) in which we substituted the Word Identification (WID) subtest of the WRMT-R (Woodcock, 1998) for WIF. With this substitution, model fit was significantly better than for WID alone, $\chi^2(4, N = 134) = 23.71$, $p < .0001$, producing a lower rate of false positives: 31 for WID with the cognitive predictors versus 59 for WID alone. This result is similar to the improvement obtained by the addition of the cognitive variables to the December WIF intercept in Model 2. With WIF in the model, RAN, phonological processing, and nonverbal reasoning were uniquely predictive; when WID replaced WIF, RAN and nonverbal reasoning were uniquely predictive, whereas phonological processing was not ($p = .076$). Thus, although WID competed somewhat more successfully with the cognitive predictors than WIF, the pattern of findings was similar for the two word-reading measures.

A fourth constraint on the generalizability of these findings was our arbitrary choice of cognitive dimensions. Although these abilities—RAN, phonological processing, oral language comprehension, and nonverbal reasoning—were selected on the basis of extant research, we could have chosen different and maybe more important dimensions such as working memory or sustained attention to model long-term disability classification. Future research could take a more comprehensive look at relations between cognitive processes and reading comprehension and RD.

Fifth, we have used adjectives like *important* and *noteworthy* in this article to characterize how well the cognitive dimensions (with and without WIF data) predicted fifth-grade RD, especially given the restricted nature of the study sample. We've not yet discussed, however, what these *important* and *noteworthy* relations might say about our sample's first-though-fifth-grade education. That first-grade cognitive characteristics and reading performance so strongly forecasted later reading comprehension suggests a lack of effective instruction in the interim. We say this because more effective instruction would likely have "disrupted" the initial rank ordering of students and compromised (weakened) the prediction model. The larger point is that the relative accuracy of predicted performance must often be understood in context: Stronger and weaker instruction will affect the magnitude of relations between cognitive and academic predictors on one hand and long-term performance and disability status on the other.

## On the Merits of Stage 2 Screening

Readers should be mindful, too, of the size and cost of our fall-of-first-grade, Stage 2 test battery. To assess RAN, we used four subtests from the CTOPP (Wagner et al., 1999). Phonological processing was based on six additional CTOPP subtests. Comprehension of oral language was derived from the Listening Comprehension component of the WDRB (Woodcock, 1997), the Vocabulary and Similarities subtests from the WASI (Psychological Corporation, 1999), and the Oral Vocabulary subtest of the WJ III (Woodcock et al., 2001). Nonverbal reasoning was based on scores on WASI Block Design and WASI Matrix Reasoning (Psychological Corporation, 1999).

Our first-grade battery was administered individually to students and lasted 110 min. The lengthy battery permitted us to group a large number of measures into conceptually related dimensions (e.g., "phonological processing") and then to conduct confirmatory factor analyses of students' scores. Findings from these analyses strengthened confidence in their validity.

As heuristic, we trimmed the battery of cognitive predictors, entering only one measure per cognitive dimension to predict RD membership in fifth grade. The measure chosen to represent each cognitive dimension had the highest correlation with the dimensional score. The utility of the trimmed battery (i.e., CTOPP RLN for RAN, CTOPP Blending for phonological processing, WDRB Listening Comprehension for oral language comprehension, and WASI Matrix Reasoning for nonverbal reasoning) was promising.

Performance on each of these four tests contributed significantly to the prediction of RD/NRD; the rate of false positives was the same ($N = 32$) for the trimmed and full battery, and testing time was reduced from 110 min to 30 min. (We could not assess whether the model fits differed because each incorporated four predictors.)

In weighing the scientific and practical importance of a lengthy first-grade, Stage 2 test battery, researchers and practitioners should recognize that, had we followed the practice of many school districts and relied on a single-stage universal screen, we would have tutored 195 students, of whom only 33 were later found to be RD. In other words, we would have tutored 162 false positives, or children who did not require the tutoring. Consider this in terms of dollars and cents. Assume the 195 children were tutored in triads (as in our study) and for 20 weeks, 3 sessions per week, and 40 min per session, totaling 2,400 min, or 40 hr. Assume further that the expense of tutoring to a district is $100 per hour. At this rate, the district is spending $4,000 to tutor a group of three students for 40 hr, or $1,333 per student. If we multiply this per-pupil cost by the 195 students identified by our Stage 1 screen, the total cost is $259,935.

Use of a Stage 2 screen reduced the number of students in need of tutoring by two thirds (from 195 to 65, 32 of whom were false positives). In considering its price, let’s say a district uses something like our trimmed, 30-min test battery with the 195 students identified by the Stage 1 assessment. This testing would require a total of 5,850 min, or 97.5 hr. If testing were conducted at $100 per hour, a Stage 2 battery would cost $9,800. To this we add the expense of tutoring the 65 children identified by the Stage 2 battery, or $86,645 ($1,333 × 65 students). Thus, the total price of a Stage 2 screen, including the subsequent tutoring, is $96,445, which is $163,490 less than the cost of the single- stage screen (i.e., $259,935 minus $96,445).

We have tried to be explicit in costing out the two-stage screen to help readers understand our calculations. They are partly based on a trimmed, Stage 2 battery that requires more validation. In addition, the nature, frequency, duration, and cost of these procedures will surely vary from place to place. Nevertheless, findings from this study indicate that a two-stage screening procedure can save school districts money.

## How to Instruct Low Responders?

We started this article by discussing the 5% or so of the K–12 student population who are not responsive to researchers’ best-evidence instruction (a proportion, we suggested, that increases substantially when practitioners do the instructing). The 5% problem has become increasingly evident in well-conducted, intervention-oriented, multitiered RTI research. We believe the 5% figure signifies not just that millions of children are not learning but that a skills-based approach with its well-known characteristics of explicitness, directness, and systematicity, which is used by many RTI researchers including ourselves, has its limits. Saying this is not to make light of such instruction. It benefits many children, but not all, which begs the question, “If not a skills-based approach, then what?”

There has been much interest, historically and currently, in cognitively focused instruction as an alternative to skills-based approaches (see Learning Disabilities Association, 2010). This enthusiasm is matched by the profound skepticism of others (e.g., Consortium for Evidence-Based Early Intervention Practices, 2010)—and for good reason. Programs to train and strengthen cognitive abilities, in hopes that such training would accelerate academic growth, have proved disappointing. Research in the 1970s and 1980s on Diagnostic Prescriptive Instruction (see Arter & Jenkins, 1979; Hammill & Larsen, 1974; Kavale, 1982) and on more recent iterations of ability training (see Kearns & Fuchs, in press, for a review) has failed to demonstrate a value added.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

However, the training of cognitive abilities is not the only way to think about the importance of cognition to instruction. A different approach is to explore whether cognitive characteristics may moderate instruction such that students with cognitive Characteristic A improve on average more so than students with Characteristic B in the same skills-based reading program; or whether students with cognitive Characteristic A generally outperform those with Characteristic B in one program but the reverse is obtained in a second program. Might cognitive characteristics, in short, cause differential responses to the same or different instructional programs? Put differently, do attributes of cognition interact with features of instruction? (See Baron and Kenny [1986] and Frazier, Tix, and Barron [2004] on the importance of moderators to better understand for whom a treatment is important and under what conditions.)

If so, it would seem reasonable to modify the instructional program to reflect the difference(s) observed in cognitive characteristics. If, for example, greater or lesser sustained attention is found to moderate students' responsiveness to a given program, then perhaps changing the schedule of treatment implementation (more frequent sessions of shorter duration) or changing its format or materials or presentation to add novelty may be useful avenues to pursue. The general point is that cognitive moderators may be potentially important not because they can become targets of remediation (a la abilities training) but because they suggest ways to tailor instruction for those not benefitting from it in its current form. Implicit is the suggestion that skills-based and cognitively focused approaches are not mutually exclusive; researchers and practitioners may be able to use both to develop more effective programs for a greater number of children with severe learning problems and to design more refined programs of research.

As described, our first-grade sample's cognitive characteristics contributed much to the prediction of their fifth-grade reading comprehension and disability status. This suggests that cognitive processes may moderate instruction. But we recognize that we have not demonstrated them to be moderators, not least because our study was not designed for such a purpose and we had no knowledge of, let alone control over, the education in which our student sample participated from Grades 1 through 5. Much work remains in this potentially important area.

## Acknowledgments

We thank Naomi Zigmond, Tom Scruggs, and Sharon Vaughn for their feedback on earlier drafts.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: We report data collected as part of the work of the National Research Center on Learning Disabilities, which was funded by the Office of Special Education Program in the U.S. Department of Education (Award Number H324U010004). Further work on this article was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development (Award Numbers HD056109 and HD15052). Views expressed are our own, not necessarily the agencies'.

## References

American Psychological Association. APA letter to Assistant Secretary Hager regarding the Individuals with Disabilities Education Act (IDEA), P. L. Washington, DC: Author; 2005 Sep 6. p. 108-446.Retrieved from http://www.apa.org/ppo/disability/ltr905.html

Arter JA, Jenkins JR. Differential diagnosis—prescriptive teaching: A critical appraisal. Review of Education Research. 1979; 49:517–555.

Badian NA. Reading disability defined as a discrepancy between listening and reading comprehension: A longitudinal study of stability, gender differences, and prevalence. Journal of Learning Disabilities. 1999; 32:138–148. [PubMed: 15499714]

Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology. 1986; 51:1173–1182. [PubMed: 3806354]

Bentler PM. Comparative fix indexes in structural models. Psychological Bulletin. 1990; 107:238–246. [PubMed: 2320703]

Castles A, Coltheart M. Is there a causal link from phonological awareness to success in learning to read? Cognition. 2004; 91:77–111. [PubMed: 14711492]

Catts HW, Fey ME, Zhang XY, Tomblin JB. Estimating the risk of future reading difficulties in kinder-garten children: A research-based model and its clinical implementation. Language, Speech, and Hearing Services in Schools. 2001; 32:38–50.

Catts HW, Gillespie M, Leonard LB, Kail RV, Miller CA. The role of speed of processing, rapid naming, and phonological awareness in reading achievement. Journal of Learning Disabilities. 2002; 35:510–525.

Catts, HW.; Hogan, TP.; Adlof, SM. Developing changes in reading and reading disabilities. In: Catts, HW.; Kahmi, AG., editors. The connections between language and reading disabilities. Mahwah, NJ: Lawrence Erlbaum; 2005. p. 50-71.

Chall, JS.; Jacobs, VA.; Baldwin, LE. The reading crisis: Why poor children fall behind. Cambridge, MA: Harvard University Press; 1990.

Compton DL, De Fries JC, Olson RK. Are RAN- and phonological awareness-deficits additive in children with reading disabilities? Dyslexia. 2001; 7:125–149. [PubMed: 11765981]

Compton DL, Fuchs D, Fuchs LS, Bouton B, Gilbert JK, Cho E, Crouch RC. Selecting at-risk readers in first grade for early intervention: Eliminating false positives and exploring the promise of a two-stage screening process. Journal of Educational Psychology. 2010; 102:327–340. [PubMed: 20689725]

Compton DL, Fuchs D, Fuchs LS, Elleman AM, Gilbert JK. Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. Learning and Individual Differences. 2008; 18:329–337.

Compton DL, Fuchs LS, Fuchs D, Lambert W, Hamlett C. The cognitive and academic profiles of reading and mathematics learning disabilities. Journal of Learning Disabilities. (in press).

Compton DL, Gilbert JK, Jenkins JR, Fuchs D, Fuchs LS, Cho E, Bouton B. Fast-tracking chronically unresponsive children into tier 3 instruction: What level of data is necessary to ensure selection accuracy? Journal of Learning Disabilities. (in press).

Consortium for Evidence-Based Early Intervention Practices. A response to the Learning Disabilities Association of American (LDA) white paper on specific learning disabilities (SLD) identification. Author; 2010 Dec.

Council for Exceptional Children. Position on response to intervention (RTI): The unique role of special education and special educators. Arlington, VA: Author; 2007 Oct. Retrieved from http://www.cec.sped.org/AM/Template.cfm?Section=Home&CONTENTID=9237&TEMPLATE=/CM/ContentDisplay.cfm

Cutting LE, Scarborough HS. Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. Scientific Studies of Reading. 2006; 10:277–299.

Dawes RM. The robust beauty of improper linear models in decision making. American Psychologist. 1979; 34:571–582.

de Jong PF. Working memory deficits of reading disabled children. Journal of Experimental Child Psychology. 1998; 70:75–96. [PubMed: 9729450]

Dickinson DK, McCabe A, Anastasopoulos L, Peisner F, Ellen S, Poe MD. The comprehensive language approach to early literacy: The interrelationship among vocabulary, phonological sensitivity, and print knowledge among preschool-age children. Journal of Educational Psychology. 2003; 95:465–481.

Ehri, LC. Grapheme-phoneme knowledge is essential for learning to read words in English. In: Metsala, JL.; Ehri, LC., editors. Word recognition in beginning literacy. Mahwah, NJ: Lawrence Erlbaum; 1998. p. 3-40.

Fletcher JM. Memory for verbal and nonverbal stimuli in learning disability subgroups: Analysis by selective reminding. Journal of Experimental Child Psychology. 1985; 40:244–259. [PubMed: 4045379]

Francis, DJ.; Fletcher, JM.; Catts, HW.; Tomblin, TJ. Dimensions affecting the assessment of reading comprehension. In: Stahl, SA.; Paris, SG., editors. Children's reading comprehension and assessment. Mahwah, NJ: Lawrence Erlbaum; 2005. p. 369-394.

Frazier PA, Tix AP, Barron KE. Testing moderator and mediator effects in counseling psychology research. Journal of Counseling Psychology. 2004; 51:115–134.

Fuchs D, Compton DL, Fuchs LS, Bouton B, Caffrey E. The construct and predictive validity of a dynamic assessment of young children learning to read: Implications for RTI frameworks. Journal of Learning Disabilities. 2011; 44:339–347. [PubMed: 21685347]

Fuchs D, Compton DL, Fuchs LS, Davis GC. Making "secondary intervention" work in a three-tier responsiveness-to-intervention model: Findings from the first-grade longitudinal study at the National Research Center on Learning Disabilities. Reading and Writing: A Contemporary Journal. 2008; 21:413–436.

Fuchs D, Fuchs LS. Introduction to response to intervention: What, why, and how valid is it. Reading Research Quarterly. 2006; 41:92–99.

Fuchs D, Fuchs LS, Harris AH, Roberts PH. Bridging the research-to-practice gap with Mainstream Assistance Teams: A cautionary tale. School Psychology Quarterly. 1996; 11:244–266.

Fuchs LS, Compton DC, Fuchs D, Hollenbeck K, Hamlett CL, Seethaler PM. Two-stage screening for math word-problem difficulty using dynamic assessment of algebraic learning. Journal of Learning Disabilities. (in press).

Fuchs LS, Fuchs D, Compton DL. Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. Exceptional Children. 2004; 71:7–21.

Gough, PB.; Hoover, WA.; Peterson, CL. Some observations on a simple view of reading. In: Cornoldi, C.; Oakhill, J., editors. Reading comprehension difficulties: Processes and interventions. Mahwah, NJ: Lawrence Erlbaum; 1996. p. 1-13.

Hammill DD, Larsen SC. The effectiveness of psycholinguistic training. Exceptional Children. 1974; 41:5–14.

Hatcher, L. A step-by-step approach to using the SAS system for factor analysis and structural equation modeling. Cary, NC: SAS Institute; 1994.

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

Hulme C, Snowling MJ, Caravolas M, Carroll J. Phonological skills are (probably) one cause of success in learning to read: A comment on Castles and Coltheart. Scientific Studies of Reading. 2005; 9:351–365.

Hulslander J, Olson RK, Willcutt EG, Wadsworth SJ. Longitudinal stability of reading-related skills and their prediction of reading development. Scientific Studies of Reading. 2010; 14(2):111–136. [PubMed: 20563241]

Kavale K. Meta-analysis of the relationship between visual perceptual skills and reading achievement. Journal of Learning Disabilities. 1982; 15(1):42–51. [PubMed: 7069285]

Kearns D, Fuchs D. Does cognitively focused instruction improve the academic performance of low-achieving students? Exceptional Children. (in press).

Keenan JM, Betjemann RS, Olson RK. Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. Scientific Studies of Reading. 2008; 12(3):281–300.

Leach JM, Scarborough HS, Rescorla L. Late-emerging reading disabilities. Journal of Educational Psychology. 2003; 95:211–224.

Learning Disabilities Association. The Learning Disabilities Association of America's white paper on evaluation, identification, and eligibility criteria for students with specific learning disabilities. Author; 2010 Feb.

Lipka O, Lesaux NK, Siegel LS. Retrospective analyses of the reading development of a group of Grade 4 disabled readers: Risk status and profiles over 5 years. Journal of Learning Disabilities. 2006; 39:364–378. [PubMed: 16895160]

Manis FR, Doi LM, Bhadha B. Naming speed, phonological awareness, and orthographic knowledge in second graders. Journal of Learning Disabilities. 2000; 33:325–333. 374. [PubMed: 15493095]

Muter V, Hulme C, Snowling MJ, Stevenson J. Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. Developmental Psychology. 2004; 40:665–681. [PubMed: 15355157]

Nation K, Marshall CM, Snowling MJ. Phonological and semantic contributions to children's picture naming skill: Evidence from children with developmental reading disorders. Language and Cognitive Processes. 2001; 16:241–259.

Nation K, Snowling MJ. Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. Child Development. 1998; 69:996–1011. [PubMed: 9768483]

Nation K, Snowling MJ. Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. Cognition. 1999; 70:1–13. [PubMed: 10193054]

National Joint Committee on Learning Disabilities. Responsiveness to intervention and learning disabilities. 2005 Retrieved from http://www.ldonline.org/about/partners/njcld#reports.

Oakhill JV, Cain K, Bryant PE. The dissociation of word reading and text comprehension: Evidence from component skills. Language and Cognitive Processes. 2003; 18:443–468.

Psychological Corporation. Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: Harcourt Brace; 1999.

Rack JP, Snowling MJ, Olson RK. The nonword reading deficit in developmental dyslexia: A review. Reading Research Quarterly. 1992; 27:29–53.

Ramus, F.; Szenkovits, G. Understanding the nature of the phonological deficit. In: Pugh, K.; McArdle, P., editors. How children learn to read: Current issues and new directions in the integration of cognition, neurobiology, and genetics of reading and dyslexia research and practice. New York, NY: Psychology Press; 2009. p. 153-169.

Raudenbush, SW.; Bryk, AS. Hierarchical linear models: Applications and data analysis methods. 2nd ed.. Thousand Oaks, CA: Sage; 2002.

Rieben, L.; Perfetti, CA. Learning to read: Basic research and its implications. Hillsdale, NJ: Lawrence Erlbaum; 1991.

Rogosa D, Brandt D, Zimowski M. A growth curve approach to the measurement of change. Psychological Bulletin. 1982; 92:726–748.

Rubin DB. EM and beyond. Psychometrika. 1991; 56:241–254.

SAS Institute. What's new in SAS 9.2. Cary, NC: Author; 2008.

Scarborough, HS. Developmental relationships between language and reading: Reconciling a beautiful hypothesis with some ugly facts. In: Catts, HW.; Kamhi, AG., editors. The connections between language and reading disabilities. Mahwah, NJ: Lawrence Erlbaum; 2005. p. 3-24.

Schatschneider C, Carlson CD, Francis DJ, Foorman BR, Fletcher JM. Relationship of rapid automatized naming and phonological awareness in early reading development: Implications for the double-deficit hypothesis. Journal of Learning Disabilities. 2002; 35:245–256. [PubMed: 15493321]

Seethaler PM, Fuchs LS. The predictive utility of kindergarten screening for math difficulty: How, when, and with respect to what outcomes should it occur? Exceptional Children. 2010; 77:37–60.

Sénéchal, M.; Ouellette, G.; Rodney, D. The misunderstood giant: The predictive role of early vocabulary to future reading. In: Dickinson, DK.; Neuman, SB., editors. Handbook of early literacy. Vol. vol. 2. New York, NY: Guilford; 2006. p. 173-185.

Singer, JD.; Willett, JB. Applied longitudinal data analysis: Modeling change and event occurrence. New York, NY: Oxford University Press; 2003.

Swanson HL, Howell M. Working memory, short-term memory, and speech rate as predictors of children's reading performance at different ages. Journal of Educational Psychology. 2001; 93:720–734.

Swets JA. The science of choosing the right decision threshold in high-stakes diagnostics. American Psychologist. 1992; 47:522–532. [PubMed: 1595983]

Tabachnick, BG.; Fidell, LS. Using multivariate statistics. 5th ed.. Boston, MA: Allyn & Bacon; 2007.

Torgesen JK. Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. Learning Disabilities Research & Practice. 2000; 15:55–64.

Vellutino FR, Fletcher JM, Snowling MJ, Scanlon DM. Specific reading disability (dyslexia): What have we learned in the past four decades? Journal of Child Psychology and Psychiatry. 2004; 45:2–40. [PubMed: 14959801]

Vukovic RK, Siegel LS. The double-deficit hypothesis: A comprehensive analysis of the evidence. Journal of Learning Disabilities. 2006; 39:25–47. [PubMed: 16512081]

Wagner, RK.; Torgesen, JK.; Rashotte, CA. Comprehensive Test of Phonological Processing. Austin, TX: PRO-ED; 1999.

Wanzek J, Vaughn S. Students demonstrating persistent low response to reading intervention: Three case studies. Learning Disabilities Research & Practice. 2009; 24:151–163.

Wolf M, Bowers PG. The double-deficit hypothesis for the developmental dyslexias. Journal of Educational Psychology. 1999; 91:415–438.

Woodcock, RW. Woodcock Diagnostic Reading Battery. Itasca, IL: Riverside; 1997.

Woodcock, RW. Woodcock Reading Mastery Tests–Revised. Circle Pines, MN: American Guidance Service; 1998.

Woodcock, RW.; McGrew, KS.; Mather, N. Woodcock–Johnson III Tests of Cognitive Abilities. Itasca, IL: Riverside; 2001.

Zhu, J. WASI manual. San Antonio: Psychological Corporation; 1999.

**Table 1**

Demographic and Performance Data for Full Sample (*N* = 195) and by RD (*n* = 36) and NRD (*n* = 98) Status

| | Full Sample | | | | | | NRD | | | | | | RD | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Raw | | Standard | | | | Raw | | Standard | | | | Raw | | Standard | |
| **Variable** | ***n*** | **%** | ***M*** | ***SD*** | ***M*** | ***SD*** | ***n*** | **%** | ***M*** | ***SD*** | ***M*** | ***SD*** | ***n*** | **%** | ***M*** | ***SD*** | ***M*** | ***SD*** |
| **Demographics** | | | | | | | | | | | | | | | | | | |
| Male | 109 | 55.9 | | | | | 53 | 54.1 | | | | | 23 | 63.9 | | | | |
| Race: African Am | 63 | 32.3 | | | | | 22 | 22.5 | | | | | 18 | 50.0 | | | | |
| Caucasian | 102 | 52.3 | | | | | 61 | 62.2 | | | | | 11 | 30.6 | | | | |
| Hispanic | 30 | 15.4 | | | | | 15 | 15.3 | | | | | 7 | 19.4 | | | | |
| ELL | 10 | 5.1 | | | | | 6 | 6.1 | | | | | 1 | 0.3 | | | | |
| Subsidized lunch | 88 | 45.1 | | | | | 34 | 34.7 | | | | | 29 | 80.6 | | | | |
| **Universal Screeners** | | | | | | | | | | | | | | | | | | |
| RLN | | | 81.33 | 25.26 | — | | | | 75.09 | 21.96 | — | | | | 94.33 | 27.27 | — | |
| WIF | | | 2.22 | 2.29 | — | | | | 2.91 | 2.67 | — | | | | 1.28 | 1.45 | — | |
| **Cognitive Predictors** | | | | | | | | | | | | | | | | | | |
| RAN (CTOPP): Rapid Letter Naming | | | 77.58 | 21.44 | 7.83 | 2.26 | | | 71.41 | 20.43 | 8.35 | 2.16 | | | 86.13 | 20.42 | 6.64 | 2.67 |
| Rapid Digit Naming | | | 69.93 | 19.09 | 7.70 | 2.12 | | | 64.62 | 17.80 | 8.40 | 1.92 | | | 78.18 | 20.35 | 6.44 | 2.26 |
| Rapid Color Naming | | | 92.35 | 27.36 | 8.41 | 2.92 | | | 87.72 | 27.82 | 9.36 | 2.79 | | | 103.22 | 26.22 | 6.92 | 2.77 |
| Rapid Object Naming | | | 107.04 | 31.15 | 8.02 | 2.76 | | | 103.37 | 35.25 | 8.58 | 3.02 | | | 110.66 | 25.70 | 7.28 | 2.48 |
| Factor | | | 0.00 | 1.00 | — | | | | −0.26 | 1.04 | — | | | | 0.41 | 0.89 | — | |
| PP (CTOPP): Nonword Repetition | | | 1.51 | 2.07 | 3.98 | 1.68 | | | 1.76 | 1.75 | 3.99 | 1.75 | | | 1.43 | 2.14 | 3.92 | 1.38 |
| Blending Words | | | 5.31 | 3.08 | 8.15 | 2.34 | | | 6.27 | 2.78 | 8.91 | 1.96 | | | 4.12 | 2.96 | 6.97 | 2.64 |
| Elision | | | 4.48 | 2.71 | 8.05 | 2.55 | | | 5.44 | 2.61 | 8.98 | 2.35 | | | 3.25 | 2.82 | 6.61 | 2.69 |
| Sound Matching | | | 10.49 | 4.94 | 8.64 | 2.08 | | | 12.87 | 4.43 | 9.58 | 1.89 | | | 7.58 | 4.55 | 7.36 | 2.44 |
| Memory for Digits | | | 9.40 | 2.75 | 8.21 | 2.74 | | | 10.06 | 2.71 | 8.86 | 2.78 | | | 8.06 | 2.84 | 6.75 | 2.79 |
| Segmenting Words | | | 5.17 | 3.76 | 7.72 | 1.98 | | | 6.32 | 3.66 | 8.31 | 1.95 | | | 2.88 | 3.24 | 6.53 | 1.70 |
| Factor | | | 0.00 | 1.00 | — | | | | 0.46 | 0.88 | — | | | | −0.67 | 0.97 | — | |
| OL: WDRB Listen Comp | | | 13.70 | 4.33 | 84.32 | 13.91 | | | 15.11 | 3.87 | 88.08 | 14.39 | | | 12.22 | 4.56 | 79.97 | 13.04 |
| WASI Vocabulary | | | 17.37 | 6.10 | 40.72 | 9.32 | | | 18.89 | 5.79 | 43.32 | 8.71 | | | 15.19 | 6.49 | 37.08 | 9.56 |
| WASI Similarities | | | 10.90 | 6.18 | 46.04 | 10.37 | | | 12.47 | 6.36 | 48.81 | 10.54 | | | 9.08 | 5.57 | 42.83 | 9.10 |

| | | | Full Sample | | | | | | NRD | | | | | | RD | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Raw | | Standard | | | | Raw | | Standard | | | | Raw | | Standard | |
| **Variable** | ***n*** | **%** | ***M*** | ***SD*** | ***M*** | ***SD*** | ***n*** | **%** | ***M*** | ***SD*** | ***M*** | ***SD*** | ***n*** | **%** | ***M*** | ***SD*** | ***M*** | ***SD*** |
| WJ III Vocabulary | | | 8.36 | 3.64 | 91.80 | 13.83 | | | 9.84 | 3.43 | 96.14 | 11.96 | | | 7.56 | 3.86 | 87.22 | 15.98 |
| Factor | | | 0.00 | 1.00 | — | | | | 0.38 | 0.96 | — | | | | −0.42 | 1.05 | — | |
| NR: WASI Matrix Reason | | | 6.55 | 3.99 | 44.26 | 7.46 | | | 8.06 | 4.28 | 47.06 | 7.60 | | | 4.50 | 2.24 | 40.22 | 5.37 |
| WASI Block Design | | | 5.54 | 3.58 | 43.98 | 7.61 | | | 6.62 | 3.72 | 46.39 | 7.22 | | | 4.39 | 3.12 | 41.25 | 7.58 |
| Factor | | | 0.00 | 1.00 | — | | | | 0.40 | 1.05 | — | | | | −0.49 | 0.70 | — | |
| **First-Grade Reading (WIF)** | | | | | | | | | | | | | | | | | | |
| December intercept | | | 24.39 | 11.43 | — | | | | 27.95 | 12.01 | — | | | | 18.36 | 10.84 | — | |
| May intercept | | | 42.92 | 21.64 | — | | | | 52.70 | 20.46 | — | | | | 24.18 | 18.41 | — | |
| **Fifth-Grade Reading (WRMT-R)** | | | | | | | | | | | | | | | | | | |
| Passage Comp intercept | | | 28.76 | 3.64 | 91.24 | 8.07 | | | 31.43 | 2.13 | 97.13 | 4.71 | | | 23.44 | 2.58 | 79.38 | 6.61 |

Note: RD and NRD students totaled 134. An additional 64 students obtained a standard score intercept on the Passage Comprehension subtest of the WRMT-R between 85 and 92. These students were in a "buffer zone." Raw is raw score. Standard is standard score ($M$ = 100, $SD$ = 15 except WASI, where $M$ = 50, $SD$ = 10, and CTOPP, where $M$ = 10, $SD$ = 3). ELL = English language learner; RLN = Rapid Letter Naming; WIF = Word Identification Fluency; RAN = rapid automatized naming; CTOPP = *Comprehensive Test of Phonological Processing*; PP = phonological processing; OL = oral language; *WDRB* = *Woodcock Diagnostic Reading Battery;* WASI = *Wechsler Abbreviated Scale of Intelligence;* NR = nonverbal reasoning; WJ III = *Woodcock–Johnson III Tests of Cognitive Abilities*; WRMT-R = *Woodcock Reading Mastery Tests-Revised*.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

**Table 2**

Correlations Among Cognitive Predictors, December WIF Intercept, May WIF Intercept, and Spring of Fifth Grade Intercept

| | Correlations | | | | | |
|---|---|---|---|---|---|---|
| **Variable** | **PP** | **OL** | **NR** | **W6** | **W18** | **5th** |
| RAN | –.15 | –.11 | –.09 | –.40 | –.47 | –.35 |
| PP | | .54 | .41 | .28 | .30 | .56 |
| OL | | | .39 | .11 | .08 | .44 |
| NR | | | | .28 | .24 | .43 |
| W6 | | | | | .83 | .51 |
| W18 | | | | | | .64 |
| 5th | | | | | | |

Note: RAN = rapid automatized naming; PP = phonological processing; OL = oral language; NR = nonverbal reasoning; W6 = WIF, December Intercept; W18 = WIF, May intercept; 5th = spring of fifth grade intercept.

## Table 3

Regression Models Predicting Individual Differences in Fifth-Grade Passage Comprehension

| | *B* | *SE* | β | *t* | *p* Value |
|---|---|---|---|---|---|
| **Using December WIF as a Predictor** | | | | | |
| Constant | 26.43 | 0.49 | | 54.14 | < .001 |
| December WIF | 0.10 | 0.02 | .30 | 5.09 | < .001 |
| RAN | –0.56 | 0.20 | –.15 | –2.76 | < .001 |
| Phonological processing | 1.10 | 0.23 | .30 | 4.77 | < .001 |
| Oral language | 0.63 | 0.23 | .17 | 2.81 | .006 |
| Nonverbal reasoning | 0.53 | 0.21 | .15 | 2.53 | .012 |
| **Using May WIF as a Predictor** | | | | | |
| Constant | 21.17 | 0.42 | | 60.25 | < .001 |
| May WIF | 0.08 | 0.01 | .50 | 9.24 | < .001 |
| RAN | –0.18 | 0.19 | –.05 | –0.97 | .335 |
| Phonological processing | 0.87 | 0.21 | .24 | 4.21 | < .001 |
| Oral language | 0.80 | 0.20 | .22 | 3.97 | < .001 |
| Nonverbal reasoning | 0.46 | 0.19 | .13 | 2.47 | .014 |

Note: WIF = Word Identification Fluency; RAN = rapid automatized naming.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

**Table 4**

Classification Indices for Predicting RD Status in Spring of Fifth Grade (*N* = 134)

| | | | | | | | | | | | | ROC | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model (1–5)** | ***B*** | ***SE*** | **Wald** | ***p*** | **TN** | **FN** | **TP** | **FP** | **Hit Rate** | **Sensitivity** | **Specificity** | **AUC** | ***SE*** |
| **1: December WIF** | –0.108 | 0.030 | 18.064 | .000 | 26 | 3 | 33 | 72 | 44.0 | 91.7 | 26.5 | .791 | .049 |
| Constant | 1.391 | 0.631 | 4.854 | .028 | | | | | | | | | |
| **2: December WIF** | –0.039 | 0.029 | 1.786 | .181 | 66 | 3 | 33 | 32 | 73.9 | 91.7 | 67.3 | .859 | .032 |
| PP | –0.089 | 0.328 | 6.105 | .013 | | | | | | | | | |
| Oral language | –0.157 | 0.287 | 0.299 | .585 | | | | | | | | | |
| Reasoning | –0.786 | 0.364 | 4.656 | .031 | | | | | | | | | |
| RAN | –0.493 | 0.250 | 3.883 | .046 | | | | | | | | | |
| Constant | –0.311 | 0.699 | 0.198 | .657 | | | | | | | | | |
| **3: May WIF** | –0.091 | 0.017 | 28.521 | .000 | 43 | 3 | 33 | 55 | 56.7 | 91.7 | 43.9 | .870 | .042 |
| Constant | 2.310 | 0.591 | 15.308 | .000 | | | | | | | | | |
| **4: May WIF** | –0.080 | 0.020 | 16.061 | .000 | 67 | 3 | 33 | 31 | 74.6 | 91.7 | 68.4 | .910 | .029 |
| PP | –0.636 | 0.357 | 3.173 | .075 | | | | | | | | | |
| Oral language | –0.378 | 0.322 | 1.382 | .240 | | | | | | | | | |
| Reasoning | –0.708 | 0.415 | 2.902 | .088 | | | | | | | | | |
| RAN | 0.139 | 0.278 | 0.250 | .617 | | | | | | | | | |
| Constant | 1.770 | 0.733 | 5.836 | .016 | | | | | | | | | |
| **5: PP** | –0.920 | 0.319 | 8.299 | .004 | 66 | 3 | 33 | 32 | 73.9 | 91.7 | 67.3 | .858 | .033 |
| Oral language | –0.101 | 0.282 | 0.128 | .721 | | | | | | | | | |
| Reasoning | –0.865 | 0.358 | 5.832 | .016 | | | | | | | | | |
| RAN | 0.616 | 0.241 | 6.547 | .011 | | | | | | | | | |
| Constant | –1.211 | 0.265 | 20.821 | .000 | | | | | | | | | |

Note: Hit Rate, Sensitivity, and Specificity are expressed as percentages. RD = reading disability; TN = true negatives; FN = false negatives; TP = true positives; FP = false positives; Hit Rate = (TP+TN)/N; Sensitivity = TP/(TP+FN); Specificity = TN/(TN+TP); ROC = receiver operating characteristic; AUC = area under the curve; SE = standard error; PP = phonological processing; RAN = rapid automatized naming;