School Psychology Review,
2011, Volume 40, No. 1, pp. 3–22

Cognitive Correlates of Inadequate Response to Reading Intervention

Jack M. Fletcher, Karla K. Stuebing, and Amy E. Barth
University of Houston

Carolyn A. Denton
University of Texas Health Science Center at Houston

Paul T. Cirino and David J. Francis
University of Houston

Sharon Vaughn
University of Texas—Austin

Abstract. The cognitive attributes of Grade 1 students who responded adequately and inadequately to a Tier 2 reading intervention were evaluated. The groups included inadequate responders based on decoding and fluency criteria ($n = 29$), only fluency criteria ($n = 75$), adequate responders ($n = 85$), and typically achieving students ($n = 69$). The cognitive measures included assessments of phonological awareness, rapid letter naming, oral language skills, processing speed, vocabulary, and nonverbal problem solving. Comparisons of all four groups identified phonological awareness as the most significant contributor to group differentiation. Measures of rapid letter naming, syntactic comprehension/working memory, and vocabulary also contributed uniquely to some comparisons of adequate and inadequate responders. In a series of regression analyses designed to evaluate the contributions of responder status to cognitive skills independently of variability in reading skills, only the model for rapid letter naming achieved statistical significance, accounting for a small (1%) increment in explained variance beyond that explained by models based only on reading levels. Altogether, these results do not suggest qualitative differences among the groups, but are consistent with a continuum of severity associated with the level of reading skills across the four groups.

A recent consensus report suggested that students with learning disabilities (LD) should be identified on the basis of inadequate treatment response, low achievement, and traditional exclusionary criteria (Bradley, Danielson, & Hallahan, 2002). The most controversial component of this report was the indication that an assessment of response to instruction is a necessary (but not sufficient) component of identification. From a classifi-

Correspondence regarding this article should be addressed to Jack M. Fletcher, Department of Psychology, University of Houston TMC Annex, 2151 W. Holcombe, Suite 222, Houston, TX 77204-5053; e-mail: jackfletcher@uh.edu

cation perspective, the validity of this provision should be tested as a hypothesis by comparing adequate and inadequate responders on attributes not used to define the groups, such as cognitive processing. If adequate and inadequate responders can be differentiated from students typically developing on these nondefinitional variables, the classification hypothesis accrues validity (Morris & Fletcher, 1988).

The consensus report excluded assessments of cognitive processing skills known to underlie different kinds of LD as a component of identification. We differentiate cognitive assessments of skills that support mental operations (e.g., language, memory, problem solving) and do not involve reading for task completion from assessments of different components of reading, such as decoding, fluency, and comprehension. The latter are also cognitive measures, but are determined in part by cognitive processes that vary with the component of reading that is assessed (Vellutino, Fletcher, Snowling, & Scanlon, 2004).

Assessing cognitive skills is controversial in school psychology because of questions about the value added by these tests for identifying or treating LD (Gresham, 2009); however, these assessments are commonly employed, and strengths and weaknesses in cognitive processes are clearly related to the achievement domains that represent LD (Reynolds & Shaywitz, 2009). Although assessment of cognitive processes is not required for identification of LD in the Individuals with Disabilities in Education Act (IDEA; U.S. Department of Education, 2004), Hale et al. (2008) proposed that inadequate responders to Tier 2 intervention should receive a cognitive assessment to explain why the students did not respond to intervention, to guide treatment planning, and as an alternative to LD eligibility models explicitly identified in IDEA (ability–achievement discrepancy and methods based on response to intervention).

This issue has significant implications for everyday practice in school psychology because it suggests a major role for cognitive assessment for intervention (and for identification). However, a recent review (Pashler, McDaniel, Rohrer, & Bjork, 2009) did not identify evidence that interventions based on group by treatment interactions (e.g., learning styles, aptitude by treatment interactions) were differentially related to outcomes. Consistent with views from other school psychologists, whether cognitive skills represent child attributes that interact with treatment outcomes and are essential components of intervention planning is not well established (Gresham, 2009; Reschly & Tilly, 1999). Moreover, little research establishes whether inadequate responders differ from adequate responders and typical achievers outside of the defining characteristics of inadequate instructional response and poor development of academic skills. Taking an approach somewhat different from the analysis of group by treatment interactions, we approached the question of cognitive assessment from a classification perspective, addressing whether there are unique cognitive attributes of inadequate responders.

Cognitive and Behavioral Attributes of Inadequate Responders

One meta-analysis has addressed whether cognitive skills represent attributes of variously defined subgroups of inadequate responders (Nelson, Benner, & Gonzalez, 2003). This meta-analysis initially utilized a literature review by Al Otaiba and Fuchs (2002), which summarized 23 studies of preschool through Grade 3 students who received reading interventions. Al Otaiba and Fuchs reported that most studies identified difficulties with phonological awareness as a major characteristic of inadequate responders. However, difficulties with rapid naming, phonological working memory, orthographic processing, and verbal skills, as well as attention and behavior problems, and demographic variables, also correlated with inadequate response.

In their meta-analysis of 30 studies, Nelson et al. (2003) began with these 23 studies. They used the same search criteria as Al Otaiba and Fuchs (2002), but disagreed on the inclusion of 4 studies and added 11 other studies. Moderate to small weighted effect sizes were reported for rapid naming (Z_r

= 0.51), problem behavior (Z_r = 0.46), phonological awareness (Z_r = 0.42), letter knowledge (Z_r = 0.35), memory (Z_r = 0.31), and IQ (Z_r = 0.26). Effect sizes for demographic and disability/retention variables were negligible. Except for the negligible weightings for demographic variables and the statistical equivalence of rapid naming, phonological awareness, and behavior problems, these results were consistent with Al Otaiba and Fuchs (2002).

Since these two syntheses, other studies have examined cognitive characteristics of students with inadequate response to reading intervention. Stage, Abbott, Jenkins, and Berninger (2003) compared cognitive attributes in students who responded “faster” or “slower” to a Grade 1 intervention. Faster responders had higher initial reading levels and reading-related language skills, including phonological and orthographic awareness, rapid naming, and verbal reasoning. Al Otaiba and Fuchs (2006) used a letter naming fluency task to classify students as consistently and inconsistently responsive to intervention in kindergarten and Grade 1. Consistently inadequate responders obtained lower scores on measures of morphology, vocabulary, rapid naming, sentence repetition, and word discrimination, and had higher rates of problem behaviors. Phonological segmentation was weakly related to responder status, with Al Otaiba and Fuchs emphasizing low verbal skills (e.g., vocabulary) as a major attribute of inadequate responders.

Vellutino, Scanlon, and Jaccard (2003) found that students who responded to Grade 1 intervention had cognitive profiles similar to typically achieving students after intervention. Before intervention, responders had been lower in phonological processing and initial levels of reading skills than typical achievers. Before and after intervention, inadequate responders were best differentiated from adequate responders on phonological awareness, rapid naming, and verbal working memory, but not verbal IQ or nonverbal processing abilities. In a second study, Vellutino, Scanlon, Small, and Fanuele (2006) used the same untimed word reading criterion (25th percentile) to classify students at the end of Grade 3 as poor readers who were difficult and less difficult to remediate. Measures of rapid naming, phonological processing, vocabulary, and verbal IQ showed a stepwise progression in accordance with the groups’ levels of word reading skills. They interpreted the relation of reading level and cognitive processing as indicating that “the cognitive abilities underlying reading ability can be placed on a continuum that determines the ease with which a child acquires functional reading skills” (Vellutino et al., 2006, p. 166).

These studies identify difficulties with phonological awareness, rapid naming, vocabulary, and oral language skills as the most consistent cognitive attributes of inadequate responders. However, these differences are relative to the samples and measures chosen for investigation, and most were ad hoc applications of responder criteria; the studies were not designed to assess differences in adequate and inadequate responders. These findings are also influenced by differences in interventions and criteria for inadequate response.

Criteria for Inadequate Response

It is difficult to specify the role of intervention differences in evaluating these studies because they vary in intensity, comprehensiveness, and grade level of the at-risk students. For the second issue, different criteria do not identify the same students as inadequate responders (Barth et al., 2008; Burns & Senesac, 2005; Fuchs, Fuchs, & Compton, 2004; Speece & Case, 2001). Fuchs and Deshler (2007) noted that methods for assessing intervention response varied both by the method and type of assessment employed. Methods include (1) final status, based on end of the year status on a criterion-or norm-referenced assessment; (2) slope discrepancy, based on criterion-referenced assessments of growth; and (3) dual discrepancy, which uses assessments of growth and the end point of the criterion-referenced assessment. Summarizing across studies, Fuchs and Deshler (2007) reported that rates of agreement were generally low when inadequate responders were identified using different methods. Another source of variability across identification methods is measurement error because of small amounts of unreliability in the identification measures and the difficulty of determining where to put the cut point on distributions that are essentially normal, a problem for any assessment or study of LD (Francis et al., 2005). Variability across identification approaches is also from differences in the types of assessments used to identify responder status, such as the use of timed assessments of word reading and passages versus untimed word reading. No studies have used a norm-referenced fluency measure with a national standardization to identify inadequate responders.

Rationale for the Present Study

Our overall research question was whether adequate and inadequate responders to a Grade 1 reading intervention can be differentiated on cognitive measures not used to define responder status, addressing the classification issue fundamental to determining whether inadequate responders might benefit from an assessment of cognitive processes. To assess differences from the type of measure employed for determining responder status, we used norm-referenced end of the year assessments of timed and untimed word reading, and performance on a criterion-referenced oral reading fluency probe. The cognitive measures included those implicated in previous studies of inadequate responders: phonological awareness, rapid naming, vocabulary, and oral language skills (Al Otaiba & Fuchs, 2006; Nelson et al., 2003). Phonological awareness, rapid letter naming skills, and vocabulary are also major correlates of poor reading (Vellutino et al., 2004). To further address oral language skills, we administered measures of syntactic comprehension and listening comprehension, both linked to reading comprehension difficulties (Catts & Hogan, 2003). In addition, we included measures of nonverbal problem solving and processing speed for broader coverage of the cognitive domain.

We hypothesized (1) regardless of reading domains (decoding vs. fluency) used to define responder status, inadequate responders will have poorer performance on measures of verbal skills (e.g., vocabulary and oral language) than adequate responders or typically achieving students; (2) phonological awareness skills will be more strongly associated with responder status when defined by decoding criteria, whereas rapid letter naming skills will be more associated with responder status when defined by fluency criteria; and (3) differences in cognitive skills between adequate and inadequate responders will reflect differences in the severity of reading impairment (i.e., a continuum of severity).

Method

Participants

The study was approved by the Institutional Review Boards at the University of Houston and University of Texas—Austin. We derived the sample from the entire Grade 1 general education population in nine elementary schools located in two study sites, one in a large urban area and the other in a smaller suburban community. These students, largely minority and economically disadvantaged, were the basis for a Tier 2 reading intervention study (see Denton, Cirino et al., in press, for the complete report). We excluded from screening only students who received their primary reading instruction outside of the regular general education classroom or in a language other than English, and those with school-identified severe intellectual or behavioral disabilities.

Figure 1 presents a flow chart illustrating student assignments to intervention groups. Denton, Cirino et al. (in press) screened 680 students for reading problems in September, identifying 461 as at risk for reading difficulties and 219 as not at risk. The large proportion of at-risk students reflects the participation of schools with many at-risk students and the use of a screening plan designed to minimize false negative errors (i.e., failure to identify a "true" at-risk child), which carries with it a higher false-positive rate (i.e., identification of students as at risk who actually develop as typical readers; Fletcher et al., 2002). Because of the potentially high false-positive

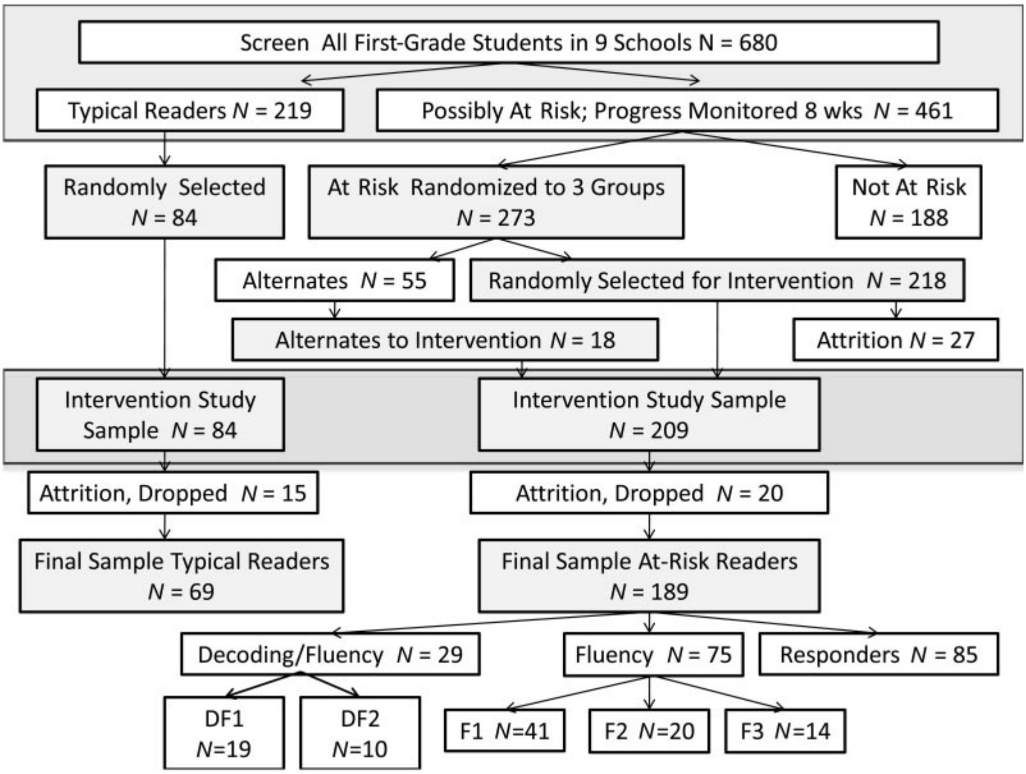

Figure 1. Flow chart showing origins of the sample for this study. DF1 = impaired on Basic Reading, TOWRE, and CMERS; DF2 = impaired on Basic Reading and TOWRE; F1 = impaired on TOWRE and CMERS; F2 = impaired on TOWRE; F3 = impaired on CMERS; Basic Reading = Woodcock-Johnson III composite of Word Identification and Word Attack (untimed decoding); TOWRE = Test of Word Reading Efficiency (timed decoding); CMERS = Continuous Monitoring of Early Reading Skills (passage fluency).

rate, students identified as at risk were then progress monitored with oral reading fluency probes from the Continuous Monitoring of Early Reading Skills (CMERS; Mathes & Torgesen, 2008) biweekly through October and November. Of the 461 identified in the initial screening, 273 failed to attain fluency benchmarks by the end of November. At this point, these 273 students were randomly assigned to one of three Tier 2 treatment groups that varied in intensity (8 weeks of instruction 2 and 4 times weekly; 16 weeks of instruction 2 times weekly).

The final sample for the current study included 189 at-risk readers who completed the intervention and the post-test assessments and 69 students identified as not at risk at the beginning of the year. The 84 at-risk readers who were initially randomized and not included in the current study were 37 students in an alternate group who did not receive treatment because of insufficient resources and another 41 who were not treated because they moved, were withdrawn by the school or parents, or did not receive sufficient intervention. Two students received intervention, but were missing post-test data, and one student with decoding, but not fluency deficits, could not be classified using our criteria. Three students were dropped because of nonverbal IQ scores <70 to exclude possible intellectual disabilities.

The 69 typically developing students were drawn from an original sample of 84 students randomly selected at the beginning of the study from the large sample of students not at risk on the screen. The 15 excluded students were 9 who moved and 6 who met inadequate responder criteria at post-test. The proportion of at-risk students assigned but not in intervention did not differ from those who remained on sociodemographic characteristics or baseline scores ($p > .05$).

Criteria for Inadequate Response

To ensure that all students who needed continued intervention were identified, we cast a broad net to identify inadequate responders, including three separately applied end-of-treatment criteria: (a) untimed word reading standard score below 91 (25th percentile) on the Woodcock-Johnson III Basic Reading Skills Cluster (WJIII; Woodcock, McGrew, & Mather, 2001); (b) word reading fluency standard score below 91 on the composite score from the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999); and (c) oral passage reading fluency score below 20 words correct per minute (wcpm on the CMERS).

The cut point for the norm-referenced tests follows previous studies of inadequate responders as well as cut points employed in many studies of LD (Torgesen, 2000; Vellutino et al., 2006). For the CMERS, the oral reading fluency criterion was selected based on the procedures used for the DIBELS, where scores <20 wcpm indicate the child is at risk (Dibels.uoregon.edu/docs/benchmarkgoals.pdfretrieved August 24, 2009). We used the DIBELS norms because there was no national sample for CMERS from which to identify a norm-referenced cut point. The cut point we used is more stringent than in other studies using CMERS (e.g., Mathes et al., 2005). The CMERS stories are less difficult than Grade 1 DIBELS stories, so equating precisely to DIBELS cut points (e.g., 25th percentile) was not justified. In support of this decision, increasing the CMERS criterion to 28 wcpm, which would represent the 25th percentile in the Hasbrouck and Tindal (2006) norms, would significantly increase the number of students identified as inadequate responders based solely on the CMERS. We show later that the increase in criterion score is difficult to justify given their scores on other reading measures. To further evaluate this cut point, we compared fluency rates in relation to the WJIII Basic Reading and TOWRE criteria. Consistent with the CMERS cut point, students in this sample who scored within 3 points of the WJIII Basic Reading criterion of 90 averaged 19.0 ($SD = 9.4$) wcpm; those with scores within 3 points of the TOWRE cut point of 90 had average CMERS rates of 24.3 ($SD = 11.5$) wcpm.

The application of these criteria yielded 85 adequate responders and 5 subgroups of inadequate responders ($n = 104$; see Figure 1). This rate may seem high, but is partly the consequence of applying multiple criteria to identifying inadequate responders and of measurement error. Figure 1 shows that the inadequate responders included two groups with primary impairments in decoding skills who *also* fell below criteria *on both fluency assessments* (DF1; $n = 19$), or who fell below criteria on the TOWRE, but not CMERS (DF2; $n = 10$). Three groups had primary problems in fluency, but not decoding. The first fluency group (F1; $n = 41$) fell below *both* TOWRE and CMERS criteria. Fluency Group 2 (F2; $n = 20$) fell below only the TOWRE criterion, whereas fluency Group 3 (F3; $n = 14$) fell below only the CMERS criterion.

Procedures

Description of the intervention

Tier 1 instruction. All students received Tier 1 instruction using evidence-based programs, a prerequisite for school selection. Beginning in September, three experienced literacy coaches held monthly meetings with Grade 1 classroom reading teachers at each school to examine graphs of their at-risk students' oral reading fluency data, discuss student progress, and provide instructional strategies. Teachers received in-class coaching on request.

Tier 2 intervention. Beginning in January, students in each intervention condition received the same Tier 2 supplemental small group reading intervention, provided by 14 trained paraprofessionals who were coached and supported by the same literacy coaches. Intervention was provided in 30-min sessions in groups of two to four students with one tutor, in a location outside of the regular classroom, on the three varying schedules previously described. The intervention was comprehensive and addressed phonemic awareness, decoding, fluency, vocabulary, and comprehension. Tutors followed a manualized curriculum with a specific scope and sequence based on a modification of the Grade 1 Read Well program (Sprick et al., 1998). Read Well is a highly structured curriculum that supports the delivery of explicit instruction in phonemic awareness and phonics, with practice in fully decodable text and repeated reading to meet fluency goals. Modifications for this study added explicit instruction in vocabulary and reading comprehension, as well as more detailed lesson scripting to support high-fidelity implementation. Read Well includes one instructional unit for each letter–sound correspondence in the scope and sequence, and four lessons are provided for each unit. Tutors used mastery tests included in the program to individualize student placement in specific units and the number of lessons taught in each unit.

Fidelity of implementation. Following procedures commonly implemented in intervention research, fidelity data were collected through direct observation of tutors three times during the semester. To document program adherence and quality of implementation, tutors were rated on Likert scales ranging from 0 (*expected but not observed*) to 3 (*observed nearly all of the time*). The mean total rating (including both fidelity and quality) was 2.47 (SD = 0.27, range 2.01 to 2.95), indicating strong implementation of the intervention (Denton, Cirino et al., in press).

Outcomes. Although there was evidence of growth in reading for many students over the intervention period, an evaluation of treatment efficacy at the end of the 8- and 16-week periods showed few differences in outcomes for the three intensity/duration groups (see Denton, Cirino et al., in press). Therefore, we combined the three groups to determine responder status.

Test administration

Students were assessed by examiners who completed an extensive assessment training program. All assessments were completed in the students' elementary schools in quiet locations. The cognitive and pretest achievement measures were administered at the end of November and December. The post-test achievement measures and nonverbal problem-solving task were administered in May. The variation in administration of cognitive assessments was from time limitations imposed by the schools. We administered the cognitive variables before intervention to ensure that they would not be influenced by treatment. The nonverbal problem-solving measure is not likely to change over the short intervention because these skills were not taught.

Measures

We selected cognitive and language measures implicated either as correlates of inadequate response or as indicators of constructs often associated with LD. The battery was necessarily parsimonious, as we were restricted to a 60-min time frame by schools, who were concerned about time lost from instruction. With some exceptions, all measures had a national standardization. A description of all tests can be found at www.texasldcenter.org/project2.asp.

Measures to determine student group membership

Woodcock-Johnson III Test of Achievement (Woodcock et al. 2001). The Basic Reading Skills composite combines the Letter-Word Identification and Word Attack subtests of untimed decoding skills. Letter-Word Identification assesses the ability to read real words; Word Attack assesses the ability to read phonetically correct nonsense words. The reliability of the composite ranges from .97 to .98 for students aged 5–8 years.

Test of Word Reading Efficiency (Torgesen et al., 1999). The Sight Word Efficiency subtest assesses the timed reading of real words presented in a list format. Phonemic Decoding Efficiency assesses timed reading of pseudowords. The TOWRE composite was the dependent variable. Alternate forms and test–retest reliability coefficients exceed .90 in this age range.

Continuous Monitoring of Early Reading Skills (Mathes & Torgesen, 2008). The CMERS is a timed measure of oral reading fluency for connected text. All texts were written at approximately a Grade 1.7 readability level according to the Flesch-Kincaid index and were 350–400 words in length. Students were required to read two passages, for 1 min each. Test–retest reliability for the first two screening periods in this study was .93. The dependent variable is the total number of words read correctly in 60 s averaged over the two stories.

Other academic measures

We gave, but did not analyze, the WJIII Passage Comprehension and Spelling subtests to permit a broader characterization of the students' academic development. Passage Comprehension uses a cloze procedure to assess sentence-level comprehension by requiring the student to read a sentence or short passage and fill in missing words. Spelling involves orally dictated words written by the student. Coefficient alpha ranges from .94 to .96 for Passage Comprehension and .88 to. 92 for Spelling in the 5- to 8-year age range.

Cognitive and linguistic measures

Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte,1999). Blending Words measures the ability to combine sounds to form whole words. Elision requires deletion of specific sounds from a word. For students 5–8 years, coefficient alpha is .96 and .99 for Elision and Blending Words subtests, respectively. To reduce variables for analysis, a composite score was created by averaging the standardized Blending Words and Elision subtests.

The Rapid Letter Naming subtest measures the speed of naming letters presented in a repeated pattern. We only administered Form A, so a standardized score could not be computed. The dependent measure was the number of letters identified divided by the total time to identify all items, and was converted into time per letter. Alternate-form and test–retest reliability coefficients are at or above .90 for students aged 5–8 years.

Clinical Evaluation of Language Fundamentals—4 (Semel et al., 2003). Concepts and Following Directions assesses the understanding and execution of oral commands containing syntactic structures that increase in length and complexity. The syntactic component involves manipulation of pronouns and sentence structure; the working memory component involves the increasing length of the commands (Tomblin & Zhang, 2006). Test–retest reliability is .87 to .88 and coefficient alpha is .90 to .92 for students 5–8 years of age. Understanding Spoken Paragraphs subtest evaluates the ability to understand oral narrative texts. The test–retest reliability is .76 to .81; coefficient alpha is .64 to .81 for students 5–8 years of age.

Underlining Test (Doehring, 1968). The Underlining Test is a paper-and-pencil measure of speed of processing (or focused attention). For each subtest, a target is displayed at the top of a page. Below are lines with the target stimuli and distracters. The participant underlines target stimuli as fast as possible for either 30 or 60 s. We used 3 subtests in which the target stimuli were (1) the number 4, nested with randomly generated single numbers; (2) a symbol (a plus sign) nested among other symbols; and (3) a diamond containing a square that also contained a diamond. The score for each subtest is the total number of correct targets identified minus errors. We computed age-adjusted residuals with a mean of 0 ($SD = 1$) for each subtest across the whole sample and then averaged these scores to create a composite.

Kaufman Brief Intelligence Test—2 (K-BIT; Kaufman & Kaufman, 2004). The Kaufman Brief Intelligence Test—2 is an individually administered intellectual screening measure. Verbal Knowledge assesses receptive vocabulary and general information (e.g., nature, geography). Matrices assesses nonverbal problem solving, requiring students to choose a diagram from among five or six choices that either "goes with" a series of other diagrams or completes an analogy. Both subtests are good indicators of overall intellectual functions. Internal consistency ranges from .86 to .89 for Verbal Knowledge and .78 to .88 for Matrices in students 5–8 years of age.

Results

Hypotheses 1 and 2 were assessed with multivariate analyses (MANOVAs) comparing group performance across the seven cognitive variables. We followed procedures in Huberty and Olejnik (2006) for a descriptive discriminant analysis that permits the interpretation of the contribution of a set of dependent variables to MANOVA. A MANOVA computes a linear composite (i.e., discriminant function) that maximally separates groups. Following Huberty and Olejnik, we used three methods for interpreting the contribution of individual variables to the discriminant function, including canonical structure correlations, standardized discriminant function coefficients, and univariate tests, where alpha per variable was set at $.05/7 = .007$. The canonical structure coefficients represented the bivariate correlation of each variable with the discriminant function maximally separating groups, whereas the standardized coefficients provided an index of the unique contribution of each variable to group separation given the set of variables selected for the model. We presented the univariate tests because there are no statistical tests associated with either of the two multivariate methods for interpretation of the canonical variates. In a comparison with two groups, the canonical structure correlations and the univariate tests will parallel one another (Huberty & Olejnik).

Although MANOVA is not affected by the scaling of the measures, visual interpretation is facilitated when the measures are on the same scale. We adjusted the raw scores of the cognitive variables for age and retained the studentized residuals, placing all of the variables into a *z*-score metric. This permitted control for the small age differences across the four groups (see Table 2). We checked each group distributions for restriction of range from either the scaling or the approach to group definitions and found no evidence of range restriction.

Comparisons of Decoding- and Fluency-Impaired Subgroups

Table 1 presents mean standard scores for the three reading measures used to define the groups, showing that the DF1 group was more impaired than the DF2 group. The F1 group, which fell below both fluency criteria, was more impaired in decoding than the F2 group, which fell below the TOWRE criterion; the least impaired group (F3) only fell below the CMERS criterion. Not surprisingly because we created the differences among groups by the cut points used to define them, analyses of variance across the five groups were significant for all three measures ($p < .0001$). Table 1 shows that the DF1 group differs from the other four groups on all three measures. The DF2 group has higher decoding (and fluency scores) than the DF1 group, but is significantly below the three F groups on decoding. The DF2 group does not differ from the F1 and F2 groups on TOWRE or from the F2 group on CMERS, consistent with the definitions. The F1 and F2 groups are similar on WJIII Basic Reading, but significantly below the F3 group, which has above average decoding scores.

To maximize sample size within groups, we evaluated whether the decoding/fluency (DF1, DF2) and reading fluency (F1, F2, F3) subgroups could be differentiated on the cog-

Table 1
Means and Standard Deviations for the Decoding Fluency and Fluency Groups on Criterion Academic Outcome Variables

	Group									
	DF1 $N = 19$		DF2 $N = 10$		F1 $N = 41$		F2 $N = 20$		F3 $N = 14$	
Task	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Basic Reading[a]	78.37	9.79	84.80	4.69	98.02	4.68	98.95	4.54	108.71	4.27
TOWRE[a]	73.79	7.68	80.00	5.31	83.83	5.02	85.85	3.28	97.36	4.70
CMERS[b]	7.95	5.19	27.25	4.89	12.95	4.90	27.83	6.47	15.50[c]	3.71

Note. DF1 = impaired on Basic Reading, TOWRE, and CMERS; DF2 = impaired on Basic Reading and TOWRE; F1 = impaired on TOWRE and CMERS; F2 = impaired on TOWRE; F3 = impaired on CMERS; Basic Reading = Woodcock-Johnson III composite of Word Identification and Word Attack (untimed decoding); TOWRE = Test of Word Reading Efficiency (timed decoding); CMERS = Continuous Monitoring of Early Reading Skills (passage fluency).
[a]$M = 100$; $SD = 15$.
[b]Words correct per minute.

nitive variables. The cognitive profiles are graphically displayed in Figure 2. For the two DF groups, the MANOVA was significant, $F(7, 21) = 2.99$, $p < .025$, $\eta^2 = .50$. Univariate comparisons were significant for CTOPP phonological awareness, $F(1, 27) = 8.94$, $p < .006$, KBIT Verbal Knowledge, $F(1, 27) = 7.24$, $p < .02$, and KBIT Matrices, $F(1, 27) = 14.50$, $p < .0007$ (all DF1 < DF2). These differences are consistent with the previously observed stepwise progression in severity of reading difficulties (DF1 < DF2) in Table 1, with Figure 2 also suggesting parallel cognitive profiles that reflect the severity of the reading problems, so we combined them for subsequent analyses. Note that the DF2 group is too small ($n = 10$) to treat separately, and the presence of decoding deficits indicates different treatment needs from the reading fluency groups; had we compared these two groups separately with the other groups, the difference would not have met a Bonferroni-adjusted critical value of alpha ($p = .05/7 = .007$).

The MANOVA of the three fluency groups across the seven cognitive variables was not significant (see Figure 2), $F(14, 132) = 1.18$, $p < .30$, $\eta^2 = .21$. No univariate comparisons achieved the critical level of alpha ($p < .05$).

Comparisons of Adequate and Inadequate Responder Groups

Sociodemographic variables. In Table 2, the frequencies for age, subsidized lunch, English as a Second Language status, and ethnicity are shown by group. There were significant differences in age across the four groups, $F(3,254) = 25.26$, $p < .0001$, with the decoding/fluency group significantly older than the other three groups, which did not differ. However, the size of the age difference is small, with a maximum of about 7 months between the decoding/fluency and adequate responder groups. This difference was addressed by residualizing for age in scaling the cognitive data. There were no significant relations of group with gender, $\chi^2(3, N = 258) = 6.80$, $p < .08$, English as a Second Language status, $\chi^2(3, N = 257) = 2.33$, $p < .51$; $p < .15$, or race, $\chi^2(12, N = 258) = 17.24$. Subsidized lunch status and group were significantly related: $\chi^2(3, N = 258) = 8.24$, $p <$

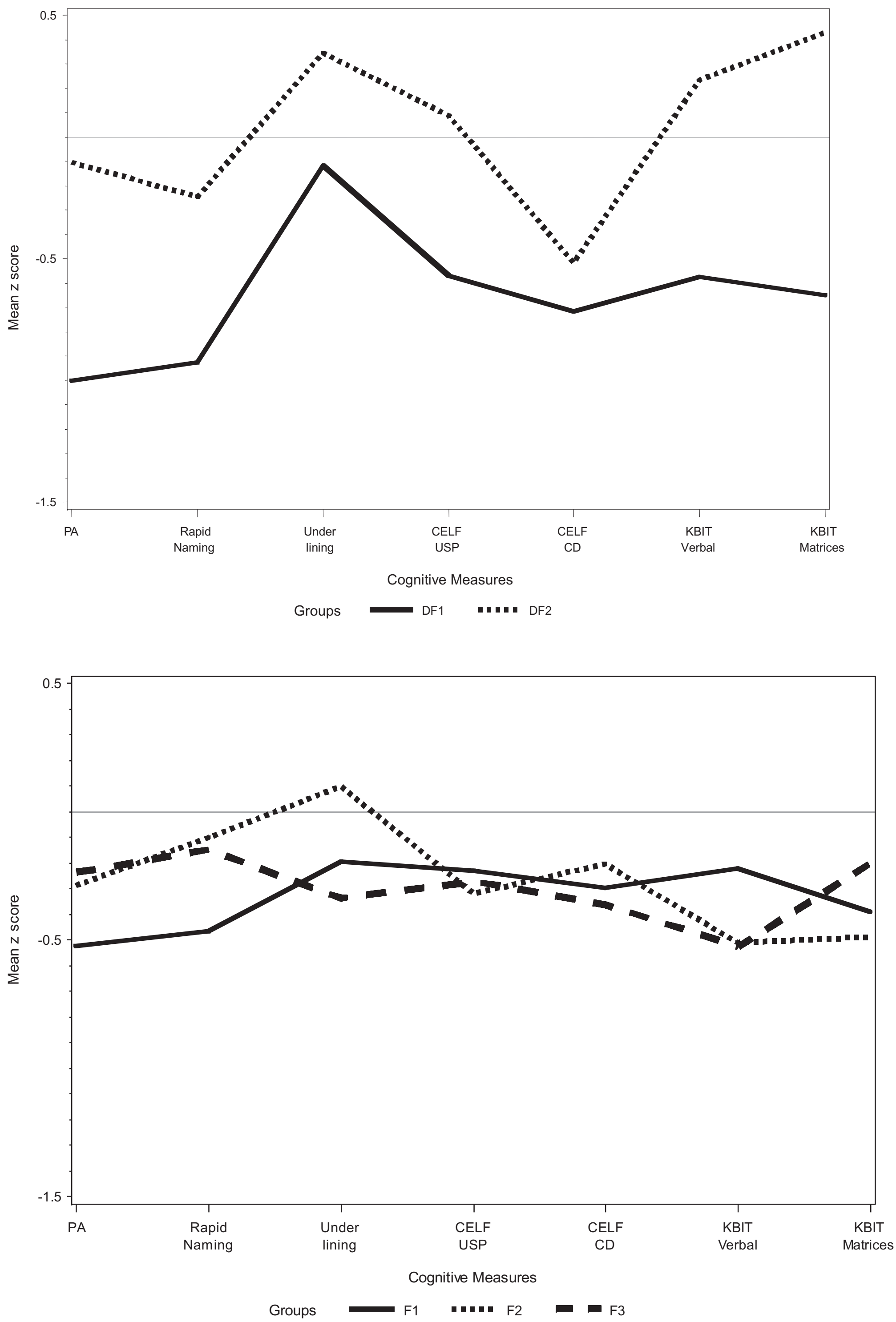

Figure 2. Cognitive profiles for inadequate responders defined by decoding and fluency (DF) criteria (upper panel) and only fluency (F) criteria (lower panel). PA = Phonological Awareness; CELF = Clinical Evaluation of Language Fundamentals—4; USP = Understanding Spoken Paragraphs; CD Concepts/Directions; KBIT = Kaufman Brief Intelligence Test—2

Table 2
Demographics by Group

	Group			
	Decoding/Fluency	Reading Fluency	Responder	Typical
Variable	$N = 29$	$N = 75$	$N = 85$	$N = 69$
Age*				
Mean	7.08	6.60	6.42	6.54
SD	0.48	0.33	0.28	0.41
% Male*	72	44	51	51
% Subsidized Lunch*	86	71	61	59
% English as Second Language	31	20	21	17
% Black	45	41	28	33
% White	7	17	19	15
% Hispanic	48	40	49	44
% Other	0	1	4	9

*$p < .05$.

.05. The participants in the decoding/fluency group were more likely to receive a subsidized lunch. Consistent with Nelson et al. (2003), this difference was small, so we did not use this variable as a covariate.

Cognitive comparisons. For Hypotheses 1 and 2, we performed a subset of all possible comparisons to control Type I error and maintain power. The decoding/fluency group was compared to the reading fluency and adequate responder groups, and the reading fluency group was compared to the adequate responder group. This permitted a direct evaluation of differences between adequate and inadequate responders defined by different reading domains. We also compared the adequate responder and typical groups to evaluate the responders' progress towards the performance levels of the not-at-risk group. The significance level was controlled at .05 by setting the alpha per comparison at .0125 (.05/4).

Table 3 provides means and standard deviations for the seven cognitive variables. A MANOVA of the age-residualized scores across the four groups was significant, $F(21, 712.67) = 9.17$, $p < .0001$, $\eta^2 = .50$; only the first discriminant function was significant ($p < .0001$). Figure 3 shows the age-residualized *z*-score profiles for the four groups. The decoding/fluency group showed the poorest performance across measures of phonological awareness, rapid naming, and syntactic comprehension/working memory (Concepts and Directions), with a stepwise progression showing higher levels of performance in the reading fluency, adequate responder, and typical groups, in that order. As Table 3 shows, this progression parallels the progression of performance on academic achievement measures used to define the groups, as well as other measures of reading comprehension and spelling. Interpretation of the significant discriminant function in Table 4 shows that phonological awareness and rapid letter naming were most strongly related to group separation.

Decoding/fluency versus reading fluency groups. The MANOVA for the decoding/fluency and reading fluency groups did not meet the critical level of alpha, $F(7, 96) = 1.89$, $p = .08$, $\eta^2 = .12$. No univariate tests met the critical level of alpha ($p < .0007$); syntactic comprehension/working memory ($p < .05$) and rapid letter naming ($p < .04$) had the largest effects.

Table 3
Performance by Group on the Cognitive and Achievement Variables in Original Measurement Units

	Group							
	Decoding/ Fluency		Reading Fluency		Adequate Responder		Typical	
	$N = 29$		$N = 75$		$N = 85$		$N = 69$	
Variable	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Cognitive variables								
Phonological Awareness	83.53	11.85	93.00	9.93	101.03	9.07	111.49	10.05
Rapid Naming	0.86	0.29	1.00	0.26	1.05	0.29	1.38	0.31
Underlining Test	0.43	0.14	0.39	0.11	0.41	0.12	0.42	0.11
CELF USP	82.24	17.81	85.07	15.82	91.08	16.45	96.99	19.03
CELF Concepts/Directions	69.31	9.98	79.07	13.65	86.59	13.98	93.62	16.80
KBIT Verbal Knowledge	79.45	13.91	82.69	16.61	91.14	14.33	95.86	16.76
KBIT Matrices	86.93	11.01	89.57	11.05	96.0	11.86	101.84	13.82
Achievement variables								
Basic Reading	80.59	8.85	100.27	6.09	111.63	8.04	118.93	10.49
TOWRE	75.93	7.48	86.89	6.82	100.77	7.09	113.57	12.11
CMERS	14.61	10.59	17.39	8.20	37.84	15.83	75.73	26.15
Passage Comprehension	76.17	9.88	88.44	5.98	98.40	6.52	107.51	9.06
Spelling	81.97	11.11	96.47	7.07	105.58	10.05	116.09	11.72

Note. Phonological Awareness = average standard score of Comprehensive Test of Phonological Processing Blending Phonemes and Elision; CELF = Clinical Evaluation of Language Fundamentals—4; USP = Understanding Spoken Paragraphs; KBIT = Kaufman Brief Intelligence Test—2; Basic Reading = Woodcock-Johnson III composite of Word Identification and Word Attack (untimed decoding); TOWRE = Test of Word Reading Efficiency (timed decoding); CMERS = Continuous Monitoring of Early Reading Skills (passage fluency). Standard scores ($M = 100$; $SD = 15$) used except for timed tests, which are targets per second.

Decoding/fluency versus adequate responder groups. The MANOVA for the decoding/fluency and adequate responder groups achieved the critical level of alpha, $F(7, 106) = 3.71$, $p < .002$, $\eta^2 = .20$. Table 4 shows that the three methods for interpreting the contribution of individual variables to the discriminant function (canonical correlation, standardized discriminant function, univariate) concurred in heavily weighting phonological awareness, rapid letter naming, and syntactic comprehension/working memory.

Reading fluency versus adequate responder groups. The MANOVA for the reading fluency and adequate responder groups achieved the critical level of alpha, $F(7, 152) = 2.86$, $p = .0008$, $\eta^2 = .11$. Table 4 shows results similar to the comparison of decoding/fluency and adequate responder groups. The three methods for interpreting variable contribution to the discriminant function concurred in heavily weighting phonological awareness. KBIT vocabulary and matrices also met the critical level of alpha for the univariate tests, but the standardized coefficients are relatively small in relation to the coefficient for phonological awareness.

Adequate responder versus typical groups. The MANOVA for the adequate responder and typical groups achieved the crit-

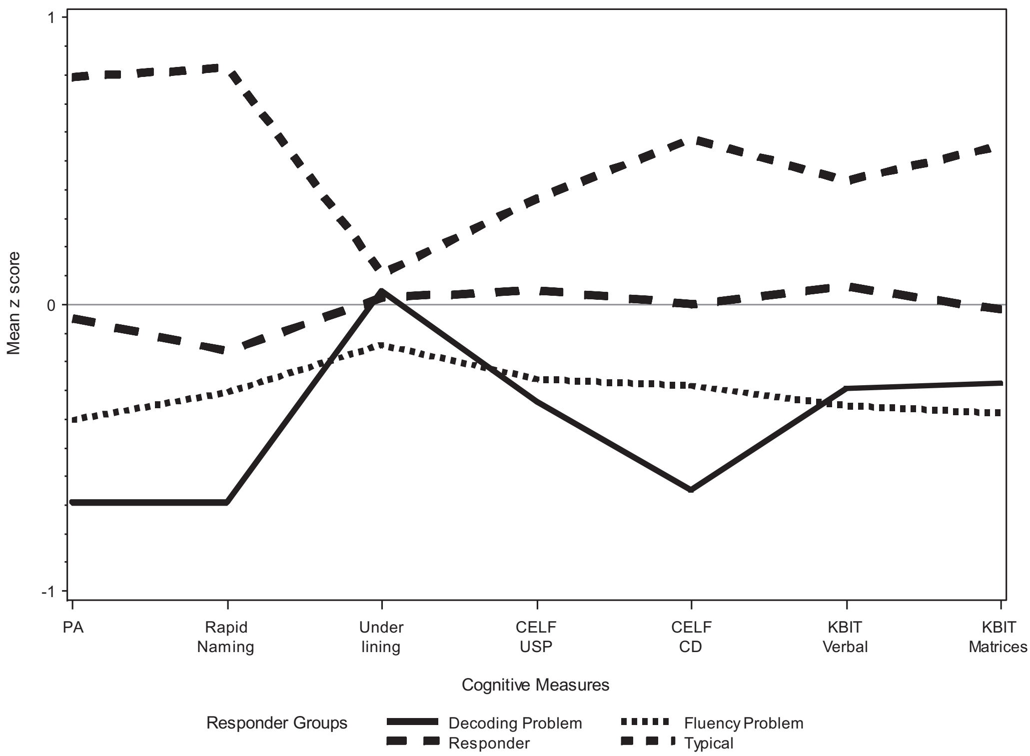

Figure 3. Mean *z* scores for cognitive measures for groups of inadequate responders who meet both decoding and fluency criteria, only fluency criteria, responders, and typical achievers. PA = Phonological Awareness; CELF = Clinical Evaluation of Language Fundamentals—4; USP = Understanding Spoken Paragraphs; CD = Concepts/Directions; KBIT = Kaufman Brief Intelligence Test—2

ical level of alpha, $F(7, 146) = 15.54$, $p < .0001$, $\eta^2 = .43$. Table 4 shows that the three methods concurred in weighting phonological awareness and rapid letter naming as the primary contributors to group separation.

Regression Analyses: A Continuum of Severity?

The test of Hypothesis 3 was based on Stanovich and Siegel (1994), who compared cognitive functions in poor readers who met and did not meet IQ–achievement discrepancy definitions. Regression models were created where each cognitive variable was predicted by the criterion-reading measures and a contrast representing, in the present study, the difference between adequate and inadequate responders. A significant beta weight for the contrast indicates variance in cognitive functions independent of reading level. This finding would suggest that a continuum of severity explanation (Vellutino et al., 2006) is an inadequate model of the relation of cognitive performance and differences between adequate and inadequate responders.

For each of the seven models, the WJIII Basic Reading, TOWRE composite, and CMERS wcpm were entered into a regression model along with a single contrast (adequate vs. inadequate responders). Before conducting the seven regressions, we investigated the suitability of the data for regression analysis by evaluating assumptions of (1) linear relations between predictors and outcome variables;

Table 4
Canonical Structure Coefficients, Standardized Discriminant Function Coefficients, and Univariate Tests by Group for Significant MANOVAs

	All Groups		DF-Responder		RF-Responder		Responder-Typical	
Variable	*r*	sdfc	*r*	sdfc	*r*	sdfc	*r*	sdfc
Phonological Awareness	.84*	.84	.85*	.71	.78*	.61	.81*	.87
Rapid Naming	.64*	.60	.54*	.41	.24	.18	.64*	.65
Underlining Test	.10	−.02	−.02	−.17	.34	.27	.06	−.07
CELF USP	.28*	−.01	.37	−.04	.48	.12	.19	.04
CELF Concepts/Directions	.46*	.09	.68*	.37	.45	−.21	.34*	−.05
KBIT Verbal Knowledge	.33*	−.05	.36	.02	.61*	.37	.23	−.04
KBIT Matrices	.40*	.26	.24	.20	.57	.38	.33	.23

Note. sdfc = Standardized discriminant function coefficient; DF = decoding/fluency; RF = reading fluency; USP = Understanding Spoken Paragraphs; PA = Phonological Awareness; CELF = Clinical Evaluation of Language Fundamentals—4; KBIT = Kaufman Brief Intelligence Test—2. The comparison of decoding/fluency and reading fluency groups did not achieve the critical level of alpha, $p < .0125$.
*Univariate $p < .007$.

(2) heteroscedasticity; (3) non-normal distribution of residuals, which may be caused by outlier data points; and (4) multicollinearity among the predictor variables (Hamilton, 1992). The evaluation of linearity suggested a quadratic term for WJIII Basic Reading in the model for rapid letter naming and for CMERS for the phonological awareness model. Heteroscedasticity was significant only for KBIT Matrices, $\chi^2(14) = 28.56$, $p < .01$, so all tests of predictor regression weights in this model used a heteroscedasticity consistent test. There was no evidence of non-normality or multicollinearity.

Across all seven models, the contrast between adequate and inadequate responders was statistically significant ($p < .05$) only for Rapid Naming, $b = -.15745$, $t(252) = -2.31$, $p < .02$. The negative sign of the b weight adjusts the predicted mean score of the responders down and the mean of the inadequate responders up relative to the prediction obtained from reading level predictors alone. The direction of the change indicated that adequate and inadequate responders are more similar on Rapid Naming than would be predicted on the basis of the reading level predictors alone. The addition of this contrast resulted in an increase in explained variance from 39% to 40%, a small increment. The group contrast did not account for unique variance in any of the other models, consistent with the hypothesized continuum of severity.

Discussion

For the overall research question, these results support the validity of classifications of LD that include evaluations of intervention response. Although we were not able to show statistically significant differences in cognitive skills between the decoding/fluency and reading fluency groups, both the decoding/fluency and reading fluency groups were clearly differentiated from the adequate responder group when separately compared (Figure 3), thus supporting the validity of the classification model proposed by Bradley et al. (2002).

We also evaluated three hypotheses concerning the cognitive attributes of adequate and inadequate responders to Grade 1 reading intervention. The first hypothesis predicted that regardless of definition, inadequate responders identified with either decoding or fluency criteria would show poorer performance than adequate responders or typically achieving students on measures of oral language. Although this hypothesis was supported in a univariate context, the support was less apparent in a multivariate context. Phonological awareness was a major contributor to group separation for the overall comparison of the four groups as well as for any significant two-group comparison. As a metacognitive assessment of language processing, phonological awareness is correlated with oral language measures, such as vocabulary, verbal reasoning, and listening comprehension, but loads on a different factor in latent variable studies in this age range (Fletcher, Lyon, Fuchs, & Barnes, 2007). Although the relation of phonological awareness and reading proficiency is well established, phonological awareness does not require reading and is often poorly developed in poor readers with strong oral language skills (Vellutino et al., 2004).

Other comparisons of general verbal skills depended on the domain of language and which groups were being compared. A measure of listening comprehension did not contribute to group separation for any comparison. The measure of syntactic processing/working memory (Concepts and Directions) contributed more robustly to comparisons of decoding/fluency and reading fluency groups versus adequate responders, but not for adequate responders versus typical students. Vocabulary did not contribute uniquely to group separation except for comparisons of the reading fluency group and adequate responders, which was also the only comparison where the nonverbal problem solving measure (KBIT Matrices) contributed uniquely. Although some univariate studies consistently identify low vocabulary/verbal reasoning as a major attribute of inadequate responders, often as a proxy for verbal intellectual capacity (Al Otaiba & Fuchs, 2006), like Stage et al. (2003) and Vellutino et al. (2006), the *unique* contribution of vocabulary to inadequate response was less apparent in a multivariate context, especially in relation to phonological awareness. Rapid letter naming skills also contributed uniquely to the separation of all four groups, the comparison of adequate responders and typical students, and the decoding/fluency and adequate responder groups.

We did not find evidence for the second hypothesis. Rapid letter naming was not more strongly related to inadequate response if fluency criteria were used. As in other studies (Vellutino et al., 2006), phonological awareness measures were stronger correlates of inadequate response than rapid letter naming measures and other language skills regardless of the reading domain used to define responder status.

The third hypothesis was largely supported. Six of the seven regression models for each cognitive variable residualized for the reading measures revealed no significant contrasts of the inadequate (combined decoding/fluency and reading fluency groups) versus the adequate responder groups. Only the contrast in the model for rapid letter naming accounted for more variance than a model with the three reading level variables. The increment in explained variance was small (1%), but demonstrated that the models were adequately powered to detect small effects. If there are unique cognitive attributes of inadequate responders, we would expect that more of these contrasts would achieve statistical significance and that effects would be larger. In fact, there was a stepwise progression similar to that observed by Vellutino et al. (2006), in which the degree of severity in the cognitive profiles paralleled the levels of reading skills across inadequate and adequate response groups (see Table 3). Because the contrast of adequate and inadequate responders was largely accounted for by the criterion reading skills, which themselves reflect a continuum of severity, these results are consistent with Vellutino et al.

Regarding the overall research question related to group differentiation on the basis of response to intervention criteria, the results indicate that a classification of LD incorporating inadequate response yields subgroups that can be differentiated on cognitive variables not used to create the subgroups. However, no single method would detect the pool of all inadequate responders. Particular concern should be expressed for the sole use of a passage reading fluency measure as in some implementations of response to intervention models. As the comparison of identification rates shows, not all students who meet norm-referenced criteria on other tests were detected with this approach. Elevating the benchmark for the passage reading fluency measure would have increased the number of students impaired only in passage reading fluency, which represents a relatively mild reading problem on a measure with lower reliability.

Limitations of the Study

The generalization of study results should be guided by our descriptions of the study sample, the intervention approach, and its implementation and outcomes. In addition, our choice of criteria for adequate intervention response should be considered. We did not incorporate criteria based on growth in this study or evaluate a dual discrepancy model based on both slope and end point assessments. We did not adequately assess verbal working memory because of the time required for these measures. We cannot determine whether syntactic comprehension versus working memory constructs account for the generally stronger contribution of the Concepts and Directions subtest to group differentiation relative to vocabulary and listening comprehension.

Group averages do not address the variability of individuals within a group, an analysis that is beyond the scope of this article. However, there are many subtyping studies based on cognitive skills that generally have not shown relations with treatment outcomes (Vellutino et al., 2004). Morris et al. (1998) identified subtypes based on profiles across eight cognitive domains that identified subtypes of poor readers with variations in phonological awareness, rapid naming, and lexical skills using the same constructs as this study operationalized via other cognitive measures, including verbal and nonverbal working memory, spatial cognition, and processing speed. Thus, it may be that the variability among individuals within the inadequate responder groups will reflect the subtypes identified by Morris et al.

The intervention from which this study was derived resulted in growth in many students. However, that study was designed to evaluate the effects of a Tier 2 intervention as commonly implemented in schools (Denton, Cirino et al., in press). However, this intervention did not generate results as robust as other early intervention studies implemented for 25 weeks or more and began in the first semester of Grade 1 (e.g., Denton, Nimon et al., 2010; Mathes et al., 2005). Findings may have been different had we delivered more intensive interventions.

The pattern of results may be different with older students, who are more likely to be impaired in reading comprehension, where the cognitive correlates are more closely associated with oral language skills (Catts & Hogan, 2003). Our study had many economically disadvantaged students. We only studied reading, so the results may not extend to LD involving math and written expression. Finally, the results do not apply to students who are assessed prior to the onset of formal reading instruction, where measures of phonological awareness, rapid naming, and vocabulary predict reading difficulties (Fletcher et al., 2007).

There are other cognitive tests that could be used, including those commonly proposed for assessing cognitive processes, such as the WJIII cognitive battery (Flanagan et al., 2007), the Cognitive Assessment System (Naglieri, 1999), and subtests from the Wechsler intelligence scales (Hale et al., 2008). Our approach focused on constructs and was limited by the amount of assessment we could complete in the context of ongoing intervention research.

Conclusions and Future Directions for Research

Studies of this sort should be completed with additional cognitive variables in the context of interventions at Tiers 2 and 3 that are more robust than the present study. In addition, there are no studies we know of that address the cognitive characteristics of inadequate responders at the secondary level. Larger studies with a mixed group of economically advantaged and disadvantaged students may be able to evaluate the heterogeneity of the inadequate responder groups and whether the two decoding/fluency groups should be combined. Other domains of LD (e.g., math, written expression) should be investigated.

The initial premise of this study was supported. Subgroups defined on the basis of inadequate response and low achievement can be differentiated on variables not used to define them. However, the differentiation seems to reflect a continuum of impairment that parallels the severity of impairment in reading skills as opposed to qualitatively distinct variation in the cognitive profiles of adequate and inadequate responders. Although more research is clearly needed, the results do not support the hypothesis of value-added benefits of assessments of cognitive processes for inadequate responders to a Tier 2 intervention (Hale et al., 2008).

The critical assessment data for reading intervention planning may be the level of reading skills and the domains of impairment (decoding, fluency, comprehension). It is noteworthy that in recent research, group by treatment interactions have been demonstrated for assessments of reading components that are directly tied to instruction. Connor et al. (2009) has shown in a series of studies that helping classroom reading teachers vary the amount of code-based versus meaning-based instruction based on strengths and weaknesses in decoding versus comprehension leads to better outcomes compared to classrooms in which this assessment information and assistance was not provided. More obviously, providing reading interventions for students with reading disabilities is more effective than providing math interventions for students with reading difficulties (and vice versa). Thus, although assessing cognitive processes for intervention purposes may not be associated with qualitatively distinct cognitive characteristics and may not justify the extensive assessments as proposed by Hale et al. (2008, in press), assessment of reading components and other academic skills appears to be well be justified.

References

Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention: A review of the literature. *Remedial and Special Education, 23,* 300–316.

Al Otaiba, S., & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. *Journal of Learning Disabilities, 39,* 414–431.

Barth, A. E., Stuebing, K. K., Anthony, J. L., Denton, C. A., Mathes, P. G., Fletcher, J. M., et al. (2008). Agreement among response to intervention criteria for identifying responder status. *Learning and Individual Differences, 18,* 296–307.

Bradley, R., Danielson, L., & Hallahan, D. P. (Eds.). (2002). *Identification of learning disabilities: Research to practice.* Mahwah, NJ: Erlbaum.

Burns, M. K., & Senesac, S. V. (2005). Comparison of dual discrepancy criteria to assess response to intervention. *Journal of School Psychology, 43*, 393–406.

Catts, H. W., & Hogan, T. P. (2003). Language basis of reading disabilities and implications for early identification and remediation. *Reading Psychology, 24,* 223–246.

Connor, C. M., Piasta, S. B., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., et al. (2009). Individualizing student instruction precisely: Effects of child by instruction interactions on first graders' literacy development. *Child Development, 80*, 77–100.

Denton, C. A., Cirino, P. T., Barth, A. E., Romain, M., Vaughn, S., Wexler, J., et al. (in press). An experimental study of scheduling and duration of "tier 2" first grade reading intervention. *Journal of Research on Educational Effectiveness*.

Denton, C. A., Nimon, K., Mathes, P. G., Swanson, E. A., Kethley, C., Kurz, T., et al. (2010). The effectiveness of a supplemental early reading intervention scaled up in multiple schools. *Exceptional Children, 76,* 394–416.

Doehring, D. G. (1968). *Patterns of impairment in specific reading disability.* Bloomington, IN: University Press.

Flanagan, D. P., Ortiz, S. O., & Alphonso, V. C. (Eds.). (2007). *Essentials of cross- battery assessment.* New York: John Wiley.

Fletcher, J. M., Foorman, B. R., Boudousquie, A. B., Barnes, M. A., Schatschneider, C., & Francis, D. J. (2002). Assessment of reading and learning disabilities: A research-based, intervention-oriented approach. *Journal of School Psychology, 40,* 27–63.

Fletcher, J. M., Lyon, G. R., Fuchs, L., & Barnes, M. A. (2007). *Learning disabilities: From identification to intervention.* New York: Guilford Press.

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 learning disabilities: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities, 38,* 98–110.

Fuchs, D., & Deshler, D. K. (2007). What we need to know about responsiveness to intervention (and shouldn't be afraid to ask). *Learning Disabilities Research & Practice, 20,* 129–136.

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2004). Identifying reading disabilities by responsiveness-to-instruction specifying measures and criteria. *Learning Disability Quarterly, 27,* 216–227.

Gresham, F. M. (2009). Using response to intervention for identification of specific learning disabilities. In A. Akin-Little, S. G. Little, M. A. Bray, & T. J. Kehl (Eds.), *Behavioral interventions in schools: Evidence-based positive strategies* (pp. 205–220). Washington, DC: American Psychological Association.

Hale, J., Alfonso, V., Berninger, V., Bracken, B., Christo, C., Clark, E., et al. (in press). Critical issues in response-to-intervention, comprehensive evaluation, and specific learning disabilities identification and intervention: An expert white paper consensus. *Learning Disability Quarterly*.

Hale, J. B., Fiorello, C. A., Miller, J. A., Wenrich, K., Teodori, A. M., & Henzel, J. (2008). WISC-IV assessment and intervention strategies for children with specific learning disabilities. In A. Prifitera, D. H. Saklofske, & L. G. Weiss (Eds.), *WISC-IV clinical assessment and intervention* (2nd ed., pp. 109–171). New York: Elsevier.

Hamilton, L. C. (1992). *Regression with graphics: A second course in applied statistics*. Belmont, CA: Wadsworth.

Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for teachers. *The Reading Teacher, 59,* 636–644.

Huberty, C. J., & Olejnik, S. (2006). *Applied discriminant analysis* (2nd ed.). New York: Wiley.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test* (2nd ed.). Minneapolis, MN: Pearson Assessment.

Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). An evaluation of two reading interventions derived from diverse models. *Reading Research Quarterly, 40,* 148–183.

Mathes, P. G., & Torgesen, J. K. (2008). *Continuous monitoring of early reading skills.* Dallas, TX: Istation.

Morris, R. D., & Fletcher, J. M. (1988). Classification in neuropsychology: A theoretical framework and research paradigm. *Journal of Clinical and Experimental Neuropsychology, 10,* 640–658.

Morris, R. D., Stuebing, K. K., Fletcher, J. M., Shaywitz, S. E., Lyon, G. R., Shankweiler, D. P., et al. (1998). Subtypes of reading disability: Variability around a phonological core. *Journal of Educational Psychology, 90,* 347–373.

Naglieri, J. A. (1999). *Essentials of CAS assessment.* New York: Wiley.

Nelson, R. J., Benner, G. J., & Gonzalez, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. *Learning Disabilities Research & Practice, 18,* 255–267.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2009). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest, 9*, 105–119.

Reschly, D. J., & Tilly, W. D. (1999). Reform trends and system design alternatives. In D. Reschly, W. Tilly, & J. Grimes (Eds.), *Special education in transition* (pp. 19–48). Longmont, CO: Sopris West.

Reynolds, C. R., & Shaywitz, S. E. (2009). Response to intervention: Ready or not? Or, from wait-to-fail to watch-them-fail. *School Psychology Quarterly, 24,* 130–145.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical evaluation of language fundamentals* (4th ed.) San Antonio TX: The Psychological Corporation.

Speece, D. L., & Case, L. P. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology, 93,* 735–749.

Sprick, M. M., Howard, L. M., & Fidanque, A. (1998). *Read Well: Critical foundations in primary reading.* Longmont, CO: Sopris West.

Stage, S. A., Abbott, R. D., Jenkins, J. R., & Berninger, V. W. (2003). Predicting response to early reading intervention from verbal IQ, reading-related language abilities, attention ratings, and verbal IQ-word reading discrepancy: Failure to validate discrepancy method. *Journal of Learning Disabilities, 36,* 24–33.

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology, 86,* 24–53.

Tomblin, J. B., & Zhang, X. (2006). The dimensionality of language ability in school-age children. *Journal of Speech, Language, and Hearing Research, 49,* 1193–1208.

Torgesen, J. K. (2000). Individual responses in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice, 15,* 55–64.

Torgesen, J. K., Wagner, R., & Rashotte, C. (1999). *Test of Word Reading Efficiency.* Austin, TX: Pro-Ed.

U.S. Department of Education. (2004). *Individuals with Disabilities Education Improvement Act*, 20 U.S.C. §1400. Washington DC: Author.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychiatry & Psychology & Allied Disciplines, 45,* 2–40.

Vellutino, F. R., Scanlon, D. M., & Jaccard, J. (2003). Toward distinguishing between cognitive and experiential deficits as primary sources of difficulty in learning to read: A two-year follow-up to difficult to remediate and readily remediated poor readers. In B. R. Foorman (Ed.), *Preventing and remediating reading difficulties* (pp. 73–120). Baltimore: York Press.

Vellutino, F. R., Scanlon, D. M., Small, S., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities, 39*, 157–169.

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

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

Date Received: September 1, 2009
Date Accepted: September 1, 2010
Action Editor: Sandra M. Chafouleas ■

Article accepted by previous Editor.

Jack M. Fletcher, Ph.D. is a Hugh Roy and Lillie Cranz Cullen Distinguished Professor of Psychology at the University of Houston. He is the Principal Investigator of the NICHD-funded Texas Center for Learning Disabilities. His research addresses classification and definition and the neuropsychological and neurobiological correlates of learning disabilities.

Karla K. Stuebing, Ph.D. is a Research Professor at the Texas Institute for Measurement, Evaluation and Statistics, Department of Psychology, University of Houston. Her research focuses on measurement issues in development disorders.

Amy E. Barth, Ph.D. is a Research Assistant Professor at the Texas Institute for Measurement, Evaluation and Statistics, Department of Psychology, University of Houston. Her research addresses the assessment of language and cognitive skills in language and learning disabilities.

Carolyn A. Denton, Ph.D. is an Associate Professor of Pediatrics in the Children's Learning Institute at the University of Texas Health Science Center at Houston. Her research is focused on interventions for children with reading disabilities and difficulties, and she is the Principal Investigator of an NICHD-funded study of interventions for children with both reading and attention difficulties.

Paul T. Cirino, Ph.D. is a developmental neuropsychologist whose interests include disorders of math and reading, executive function, and measurement. He is an associate professor in the Department of Psychology at the Texas Institute for Measurement, Evaluation and Statistics at the University of Houston.

David J. Francis, Ph.D. is a Hugh Roy and Lillie Cranz Cullen Distinguished Professor and Chairman of Psychology at the University of Houston. He is the Director of the Texas Institute for Measurement, Evaluation and Statistics with a long-term focus on measurement issues in learning disabilities.

Sharon Vaughn, Ph.D. is the H.E. Hartfelder/Southland Corp Regents Chair at the University of Texas at Austin and the Executive Director of the Meadows Center for Preventing Educational Risk. She is interested in intervention studies of a variety of populations with reading problems.