School Psychology Review,
2003, Volume 32, No. 4, pp. 557-582

The Validity of a Response-to-Instruction Paradigm to Identify Reading Disabilities: A Longitudinal Analysis of Individual Differences and Contextual Factors

Lisa Pericola Case, Deborah L. Speece, and Dawn Eddy Molloy
University of Maryland

Abstract. Using a mixed-methodology, longitudinal design, individual differences and contextual factors related to differential response to general education instruction were examined. We tested a response-to-instruction model reflecting the first three phases of a model proposed by Fuchs and Fuchs (1998). We classified first- and second-grade children at-risk for reading problems into one of three responsiveness groups and compared groups on reading, phonological processing, behavioral, and instructional context measures. Further, we examined qualitatively the interaction of individual differences and instructional context. The most persistently nonresponsive group scored significantly lower on all individual difference measures, but did not experience poorer instructional settings. Further, this group demonstrated greater difficulty learning in the general education setting. The response–to-instruction model demonstrated construct and social validity with implications for an improved prereferral system.

The field of learning disabilities (LD) is suffering an identity crisis or, perhaps more accurately, an identification crisis. An impressive body of research calls into question the use of intelligence tests as an LD marker and, by extension, achievement discrepancies derived from intelligence test scores (Fletcher et al., 1994; Gresham & Witt, 1997; Share, McGee, & Silva, 1989; Siegel, 1988; Speece & Case, 2001; Stanovich & Siegel, 1994; Stuebing et al., 2002). A recent report from the National Research Council on disproportionate minority placement in special education called for a halt to the primary use of intelligence tests in LD diagnosis (Donovan & Cross, 2002), a position aligned with the citations above. The President's Commission on Excellence in Special Education (2002) emphasized response to instruction and progress monitoring in their recommendations for assessment and identification.

The extant data base is rooted primarily in research on reading, but findings typically are generalized to the entire LD field. Three of the most frequently cited problems with intelligence tests and the discrepancy formula are (a) failure to identify children early enough to provide effective intervention services, (b) failure to find meaningful differences between children with reading achievement discrepancies and nondiscrepant poor readers, and (c) lack of instructional implications from intelligence test protocols. The criticisms of current school-identification methods bring to the forefront the need for alternative approaches. The purpose of this investigation was to examine the validity of an alternative to IQ-achievement discrepancies in identifying reading disabili-

Correspondence regarding this article should be addressed to Lisa Pericola Case, PhD, Department of Special Education, University of Maryland, Room 1308, Benjamin Building, College Park, MD 20742; E-mail: gcases@bellatlantic.net

ties in at-risk, early elementary school children. Specifically, we examined a method that relied on children's responsiveness to general education interventions. We compared children who responded differentially to this instruction on individual difference and contextual measures. Further, we conducted qualitative analyses of child, teacher, and instructional data to enhance the interpretation of the quantitative findings.

Despite the problems cited above and the attention given to reading failure in recent years (e.g., National Reading Panel, 2000; Snow, Burns, & Griffin, 1998), methods to identify children accurately remain illusive. A strong case can be made for a more seamless system of early identification, intervention, and special education services (e.g., Donovan & Cross, 2002; L. S. Fuchs, 1995; Fuchs & Fuchs, 1998; Jenkins & O'Connor, 2002). However, the measurement procedures are not well refined, resulting in missed cases (false negatives) and overidentification of children (false positives) (Jenkins & O'Connor, 2002; Scarborough, 1998; Torgesen & Burgess, 1998).

Part of the problem with traditional approaches to early identification of learning/reading disabilities is the emphasis on individual differences and the cut points used to identify the "correct" children. This focus on individual differences comes at the expense of knowledge about and contribution of the instructional environment to the child's academic growth. Definitions and classification of learning and reading disabilities emphasize intrinsic causes and exclude contextual factors, including instruction (Deno, 1989; Keogh & Speece, 1996). Duncan and Raudenbush (1999) commented that "social science research is far from definitive about whether 'context matters'" despite the fact that contexts such as families, neighborhoods, and schools "are essential to making the child fully human" (p. 29). From a different angle, Gerber (2002) criticized LD research that relied solely on research-identified samples and individual differences because it "tried to eliminate, avoid, or disregard the unexplained variance associated with the unhappy fact that special education occurred in the context of schooling" (p. 343).

The difficulties with the current system are easier to identify than the solutions. One possibility gaining attention is responsiveness to instruction (Donovan & Cross, 2002; L. S. Fuchs, 1995; Fuchs & Fuchs, 1998; MacMillan & Siperstein, 2002). Instead of measuring a child's skills at a single point in time and declaring the child as at-risk for later academic difficulty or as LD, interventions are introduced in concert with frequent measurement. Children who fail to respond to well-implemented instruction may be candidates for special education or more intensive early intervention efforts. In this way, instructional context is taken into consideration. Fuchs and Fuchs (1998) refer to this approach as "treatment validity"; other terms include "titration" or "tiered" identification (MacMillan & Siperstein, 2002). Few studies apply this approach to special education identification, but investigators using a version of response-to-instruction report favorable results (Speece & Case, 2001; Vellutino et al., 1996). Further, Speece and Shekitka (2002) asked academics in learning and reading disabilities to identify criteria to operationalize reading disabilities. Two-thirds of the 113 respondents favored response-to-instruction as a criterion.

Fuchs and Fuchs (L. S. Fuchs, 1995; Fuchs & Fuchs, 1998) offered the most fully developed conceptualization of a responsiveness-to-instruction paradigm for special education identification. As proposed by Fuchs and Fuchs (1998), children may be considered at-risk for academic problems and identified as dually discrepant (DD) if *their level of performance and rate of performance* (i.e., slope) fall below the level and rate of classmates. Given a reasonable instructional environment, in which most children are thriving, children who exhibit this dual discrepancy in level and rate of performance require intervention. Thus, children's progress may be viewed as consisting of both individual differences and instructional contexts. The response-to-instruction approach to identification requires progress monitoring (e.g., curriculum-based measurement) and well-designed, well-implemented general education interventions. Dually discrepant children who do not demonstrate progress after receiving

general education interventions may require special education services.

Speece and Case (2001) investigated the validity of the alternative identification method for reading disabilities. Based on data collected across most of the school year with an epidemiological sample of first- and second-grade children, children exhibiting dual discrepancies in reading were more impaired on phonological measures and teachers' ratings of academic competence compared to children who were either IQ-reading achievement discrepant or low achieving. Further, the DD group had equal proportions of females and males, reflected the majority/minority ethnic proportions of the population, and were younger than the other two groups of poor readers. Thus, children identified as DD represented a valid group of poor readers without prompting the negative social and political consequences associated with overrepresentation of males and minorities in special education. Further, DD children were younger, yielding implications for early identification.

Speece and Case (2001) did not include specially designed general education interventions as required by the Fuchs and Fuchs (1998) response-to-instruction model. Instead, they investigated the validity of the dual discrepancy classification within the context of general education. The purpose of the present investigation was to explore more fully the validity of an identification system that relies on response to instruction. We compared three groups of at-risk children. Groups were formed based on frequency of identification as DD across the 3 years of the study. Frequently dually discrepant (FDD) children were identified four or more times, infrequently dually discrepant (IDD) children were identified between one and three times, whereas never dually discrepant (NDD) children were not identified. Children who were identified as DD received an intervention in their general education classroom.

These groups provided the basis to pursue three research goals. The first goal was to determine if FDD children exhibited more severe reading, phonological, and behavioral problems than at-risk children who were either IDD or NDD. Children who demonstrate a consistent pattern of nonresponsiveness should be more impaired on reading and reading-related variables if the response-to-instruction method has merit. This pattern would indicate that more focused, intense interventions are required to address these children's learning needs. We also compared the groups on the number of special services provided by the school. Because our project did not supplant any of the existing programs and services in the participating school, children who were nonresponsive to general education interventions likely received additional services outside the classroom, providing an index of severity from the school's perspective and, correspondingly, social validity.

The second goal was to determine if these groups differed on several measures of instructional context. Many reasons, besides individual differences, may explain why children do not respond to general education interventions. Most salient for school psychologists and special educators are the instructional conditions in the classroom and larger school environment that may be amenable to intervention. We compared the groups of at-risk students on ratings of the classroom instructional environment, teacher experience, and classroom-level reading skill. We anticipated that children who do not respond to classroom-level instruction also may experience poorer quality classroom activities, novice teachers, and/or classmates with lower reading skills.

The third goal was to provide a qualitative analysis of both individual differences and contextual variables that may distinguish children at-risk who demonstrate a pattern of nonresponsiveness from their peers at-risk who do respond. This approach provided an opportunity to study more subtle failure-associated factors not easily captured by quantitative methods. We analyzed teacher interviews, classroom observations, and notes from intervention meetings in addition to the test data. Qualitative data were used to extend the quantitative findings and to guide thinking and further research on responsiveness to instruction.

To accomplish these goals, data based on a subsample of children at-risk from a larger investigation (Speece & Case, 2001) were ana-

lyzed. These children received specially designed general-education interventions if they were identified as DD on curriculum-based, oral-reading-fluency measures, at any point across the 3 years of the study. The interventions were developed during collaborative consultation meetings attended by the researchers, the general education teacher, and other school personnel as appropriate. Research staff monitored intervention implementation; conducted teacher interviews and classroom observations; and collected data on reading, phonological processing, intelligence, demographics, and classroom behavior. The specific research questions were (a) across 3 years, are children who are identified frequently as DD more impaired on reading, reading-related, and behavioral measures and do they experience less optimal learning contexts? and (b) what are the similarities and differences in the instructional milieu and child characteristics between children at-risk who do and do not respond to general education interventions? How do these descriptive profiles assist our understanding of the development of reading problems and reading success?

Method

Overview of the Adapted Response-to-Instruction Model

The model proposed by Fuchs and Fuchs (1998) has several key requirements: (a) frequent measurement with technically adequate measures, (b) identification of children as DD compared to classmates, and (c) faithful implementation of well-designed general education interventions. Use of curriculum-based, oral-reading-fluency measures fulfilled the first two requirements. The strong psychometric characteristics of CBM measures are well documented (e.g., Deno, 1985; Fuchs & Fuchs, 1998; Shinn, 1989). Identification of students was accomplished with The Basic Skills software program (Fuchs, Hamlett, & Fuchs, 1997). Dual discrepancies were defined as performance one standard deviation below the classroom level (mean of last two data points) *and* classroom slope (OLS regression).

To address the general education intervention requirement, we instituted collaborative consultation meetings with teachers every time a child was identified as DD. The purpose of the meetings was to develop a research-based, reasonable, and feasible classroom intervention. Meetings followed a set protocol in which we discussed the child's background, strengths, and weaknesses from the teacher's and researchers' perspectives; generated several possible interventions and their requirements; and selected one that the teacher believed she or he could implement. The interventions developed with teachers were based primarily on research-validated practices. About one-third of the interventions involved partner reading activities, another one-third involved phonological awareness and phonics instruction, and others focused on behavioral or motivational needs and reinforcement of skills through computer programs. Teachers agreed to and modified practices as their individual contexts demanded. For example, to help students develop fluency, we developed partner reading activities based on the Peer-Assisted Learning Strategies (PALS; Fuchs, Fuchs, Mathes, & Simmons, 1997; Mathes, Grek, Howard, Babyak, & Allen, 1999). As described in the literature, teachers pair each student, and student pairs complete oral repeated reading, sentence by sentence, for a specified length of time, followed by brief comprehension activities. Although this is designed for an entire class, most teachers chose to involve only target students with a higher functioning peer. Some students completed the oral reading sentence by sentence, whereas others repeated entire pages of text, depending on the teacher's evaluation of what would keep students engaged.

For each intervention, we developed a written plan that included materials needed, instructional steps, frequency and duration of instruction, group size, and motivational strategies. Teachers implemented each intervention for 8 weeks. Within the first week of implementation, research staff conducted a "support check," consisting of a scheduled observation and meeting with the teacher to ascertain that the plan could be carried out with fidelity. At this meeting, plans were adjusted as needed. Over the next 7 weeks, researchers made at

least two unannounced fidelity checks. At 4 weeks, researchers provided another support check during which the researcher met briefly with the teacher to review progress to date and to discuss any implementation difficulties. Across students and years, intervention fidelity for individual plans ranged from 0 to 1.0, with a median of .67.[1]

The model implemented was an adaptation of the Fuchs and Fuchs (1998) model because (a) we did not assess every child in each classroom but rather a subsample (explained below), (b) classrooms that had relatively low levels of performance were not first provided with a whole-class intervention, and (c) children identified as DD after two general education interventions were not given a trial intervention in special education.

The School and Reading Curriculum

The participating elementary school served 680 students from kindergarten through fifth grade and was located in a suburban, large school district in the middle-Atlantic states. The demographic features of the school population included 24.9% receiving free or reduced-price lunch with a racial distribution of 18.2% African American, .3% American Indian, 7.6% Asian, 18.5% Hispanic, and 55.3% White. The mobility rate was 16.8%.

In the first year of the study, the school district's reading curriculum followed a whole language tradition to beginning reading in that children's literature formed the primary basis of instruction with little explicit attention to decoding. As part of our work developing interventions with the teachers, we provided several inservice meetings on phonological awareness and the alphabetic principle. Teachers appeared unfamiliar with phonological awareness and the implications for instruction with young readers. In the second year of the study, the district implemented an intensive reading program in all first and second grade classrooms in Title 1 schools. This new literacy program included a 2-week, full-day summer teacher training session that emphasized phonological awareness, the alphabetic principle, writing, and assessment using running records. The implementation of the literacy program involved two certified general education teachers in each classroom for 90 minutes per day.

Participants

Participants (children and teachers) were part of a larger, 3-year investigation of three elementary schools. In this article we report on participants from the only school in which we implemented an adaptation of the Fuchs and Fuchs (1998) model for reading problems.

Across 3 years, 25 general education teachers in Grades 1 through 4 participated in the study. The average number of years of teaching experience was 9.7 (SD = 10.4). Ten teachers had a master's degree, 6 had a bachelor's degree plus 30 hours, and 9 had a bachelor's degree, all in elementary education. Ethnic representation was African American (1), Asian (1), Hispanic (2), Multiracial (2), and White (19). One teacher was male.

Two groups of first- and second-grade children were selected based on the results of screening all children on curriculum-based reading measures (CBM) in late September–early October of the study's first year. First grade children (N = 124) received two Letter Sounds Fluency probes and second-grade children (N = 127) received two Oral Reading Fluency probes. Children were deemed at-risk for reading failure if mean performance on their CBM probes placed them in the lowest 25% of their classroom. We excluded children who had an academic IEP (not including speech and language services alone) and those who, according to the teacher, did not possess enough English to benefit from reading instruction in English. The at-risk group, consisting of 53 children in Year 1, constituted the study's longitudinal sample. As a result of attrition, 36 (67.9%) of the students at-risk remained in the study for 3 years. There were no statistically significant associations between the students who left the school and the longitudinal at-risk sample on grade, X^2 (1, N = 53) = .59; race, X^2 (1, N = 53) = .12; gender, X^2 (1, N = 53) = .0001; age, F (1, 51) = .04; intelligence, F (1, 51) = .10; Letter Word Identification, F (1, 50) = .40; or mother's years of education, F (1,46) = 3.04, p = .09.

To estimate the classroom level and slope necessary to identify DD children, a second group of children, referred to as the Purposive Sample, was also selected based on the fall CBM screening data to represent the range of reading (or prereading) skill present in each classroom. Two children scoring at the class median and one each at the 30th, 75th, and 90th percentiles were selected. Thus, along with the children at-risk, five children from each classroom each year received weekly CBM assessments. In calculating level and slope, we weighted the CBM scores of the Purposive Sample children to reflect the proportion of children they represented in the classroom. These children are not the focus of this study and are not described further.

Dually Discrepant (DD) groups. A primary hypothesis of this study was that children at-risk who are frequently identified as DD should exhibit more severe reading problems than children at-risk who are less frequently or never identified as DD. To test this assertion, we formed three groups of children based on their DD status across 3 years. Continuous monitoring with CBM reading measures was the basis for determining DD status. Across 3 years, students at-risk had 9 or 10 opportunities to be identified and, thus, could move in and out of risk status across the project period. Three groups of children at-risk were formed: never dually discrepant (NDD, N = 12), infrequently dually discrepant (IDD, N = 17), and frequently dually discrepant (FDD, N = 7). The IDD children were identified three or fewer times whereas the FDD children were identified four or more times across 3 years. These criteria were selected to represent different levels of severity. FDD children identified four or more times across 3 years would have to be identified twice in a single year, consistent with Fuchs and Fuchs (1998) criteria for a trial special education placement. Therefore, we viewed the FDD children as requiring more intensive instruction. Although children were identified as dually discrepant throughout the project, classification as NDD, IDD, or FDD was not made until the end of the study. Table 1 contains demographic information for each group and total sample for Year 1.

It should be noted that in the model proposed by Fuchs and Fuchs (1998), the process of identifying children as candidates for special education conceivably could be accomplished within 1 school year after two 8-week general education interventions. The model we tested did not have special education identification as a goal because the school maintained its adherence to local, state, and federal policies governing special education placement. However, we viewed the FDD children as representative of those who would be considered for special education in the Fuchs and Fuchs model because of their persistent (across 3 years) problems in achieving the mean level and rate of progress accomplished by their peers.

Student profiles. In addition to examining group differences, another objective was to examine how individual differences and contextual factors might influence or interact with the students' reading development and responsiveness to instruction. We selected 8 participants, 4 each from the IDD and FDD groups, and analyzed the quality of reading instruction in their classrooms, teachers' background, teachers' definition and philosophies on children at-risk for reading failure, and implementation of interventions planned as a result of the students' DD status. Students were matched on grade level and gender. These pairs served as the basis for qualitative analysis of individual differences and contextual influences that may illuminate differential response to instruction. These children are described further in Results.

Measures

Child Measures

Background variables. Parents provided information on their years of education and their child's birthdate and race. For race, we used the school system's designations (African American, American Indian, Asian, Hispanic, White/Caucasian) and also provided categories for Multiracial and Other.

Reading achievement. We administered to all children three subtests from the Woodcock-Johnson Psycho-Educational Battery—Revised (WJ-R, Woodcock & Johnson,

Table 1
Demographic Information for Total Sample and Dually Discrepant Subsamples

	Sample ($N = 36$)	NDD ($N = 12$)	IDD ($N = 17$)	FDD ($N = 7$)
Grade (N)				
1	21	8	9	4
2	15	4	8	3
Gender (N)				
Female	17	5	9	3
Male	19	7	8	4
Race (N)				
Majority	11	4	5	2
Minority	25	8	12	5
Full Scale IQ				
M	92.3	99.6	89.9	85.9
SD	16.1	12.9	12.6	14.2
Mother's Years of Education				
M	12.1	11.3	11.9	13.7
SD	4.7	5.3	5.2	1.5
Age				
M	7.2	7.0	7.3	7.2
SD	.6	.6	.6	.8

Note. NDD = At-risk, not dually discrepant; IDD = At-risk, infrequently dually discrepant; FDD = At-risk, frequently dually discrepant.

1989): Letter-Word Identification, Word Attack, and Passage Comprehension. These subtests have adequate reliability for the age range of the study sample (.80–.88, Woodcock & Johnson, 1989) and adequate validity (Salvia & Ysseldyke, 1995). Age-based W scores were used in the data analysis. W scores (derived Rasch scores) are equal interval scores with a mean of 500 centered at approximately fifth grade (Woodcock & Mather, 1989, 1990).

Phonological processing battery. In the fall, subtests of an experimental version of the Comprehensive Tests of Phonological Processing (Torgesen & Wagner, 1997) were administered. The Blending and Elision subtests had 29 and 25 items, respectively, and tapped different facets of phonological awareness. The blending task required students to listen to the examiner verbalizing isolated sounds (one per second) that made up a word and then to blend these sounds into a word. Elision required the child to say a word after deleting a syllable or phoneme(s) (e.g., mat without the /m/). The Blending and Elision raw scores were summed to provide a

composite measure (Total Phonological Awareness) because of evidence that the two skills represent a single construct (Schatschneider, Francis, Foorman, Fletcher, & Mehta, 1999).

Rapid Object Naming, a type of Rapid Automized Naming (RAN) task, required students to identify five pictures randomly sequenced across four rows. The raw score was the average time, in seconds, needed to name the objects on the page on two consecutive trials. Lower scores are desirable. Word Reading Efficiency measured the number of increasingly difficult, high frequency words read correctly in 45 seconds. The measure of interest was the average number of words read correctly across two trials. All four phonological tasks have demonstrated adequate reliability (internal consistency) and predictive validity with measures of reading (Torgesen & Wagner, 1998; Wagner, Torgesen, Laughon, Simmons, & Rashotte, 1993; Wagner, Torgesen, & Rashotte, 1994).

Classroom behavior. Classroom teachers completed the Social Skills Rating System–Teacher (SSRS; Gresham & Elliott, 1990) in early spring. Teachers rated Social Skills (30 items) and Problem Behaviors (18 items) on a 3-point scale. For Academic Competence (9 items), teachers compared the target student to classmates on a 5-point scale. Scores were summed within subscale for data analysis. Higher scores on Social Skills and Academic Competence represent favorable ratings whereas higher scores on Problem Behavior represent less positive ratings. The SSRS received strong, positive evaluations on standardization, reliability (internal consistency > .80), and validity (content, criterion-related, social, and construct) by independent reviewers (Benes, 1995; Furlong & Karno, 1995).

Intelligence. Four subtests (Information, Similarities, Block Design, and Digit Span) from the Wechsler Intelligence Scale for Children—Revised (WISC-R; Wechsler, 1974) were given to all children. Full Scale IQ was estimated based on the formula provided by Sattler (1988).

Curriculum-based measures. Two fluency measures, Letter Sounds Fluency (LSF) and Oral Reading Fluency (ORF), were developed and used for initial screening in Year 1 (LSF-first grade, ORF-second grade) and ongoing progress monitoring (ORF-first through fourth grades). These were the data used to identify the three DD groups. CBM has an extensive research base documenting development procedures and psychometric characteristics (e.g., Deno, 1985; Fuchs & Fuchs, 1998; Shinn, 1989). LSF probes were developed by randomly arranging the 26 lower-case letters of the alphabet in two columns. In the standard set of directions, children were given three examples to emphasize letter sounds over letter names and to say as many sounds as possible in 1 minute. If they finished the lists before 1 minute elapsed, they were directed to begin again. The measure of interest for the first-grade screening was the number of letter sounds said correctly in 1 minute averaged across two consecutive trials.

ORF probes were developed for each grade (1–4) based on the participating school system's reading curriculum (Speece & Case, 2001). Children were provided standard directions that included a sentence summary of the passage and asked to read as many words as possible in 1 minute. Number of words read correctly in 1 minute was the measure of interest.

Alternate-forms reliability and criterion-related validity of the two measures were calculated from the screening data. The reliability coefficients were .93 and .94 for LSF and ORF, respectively. Predictive validity was determined by correlating the fall screening CBM scores with spring WJ-R Basic Reading Cluster scores for all children administered both measures. The validity coefficients were .66 for LSF ($N = 138$) and .78 for ORF ($N = 135$, Grade 2).

School attention. At the end of each year, we obtained data on the number of services each child received that year. These services represented extra-classroom attention and included prereferral meetings, special education eligibility meetings, development of an IEP, reading instruction from a reading specialist, services from a special educator on a nonhandicapped basis, English as a Second Language (ESL) instruction, and retention. The variable of interest was the sum of the number of services received.

Context Measures

Classroom slope. We considered the end-of-year slope for each classroom as one of several indicators of the quality of classroom reading instruction. In general, higher slopes are associated with stronger instructional programs.

Classroom observations. To aid in developing descriptive profiles of each classroom, research staff conducted two formal observations each year, one in the fall and one in the spring during language arts instruction. Observations ranged from 45 minutes to 2 hours, depending on the length of time teachers identified as their reading/language arts instructional period. Extensive fieldnotes were collected to capture the nature of reading instruction, instructional modifications, and classroom management. For each observation, a child was randomly chosen to help focus note taking. With the exception of the Year 1 fall observation, the child targeted for observation was identified by research staff from the pool of DD students in each classroom, or from students identified as at-risk if no DD students were present. During classroom observations, the researcher participated in classroom life only as much as necessary to maintain a comfortable environment and to create natural interactive relationships with students. The researcher informally spoke with students and teachers during noninstructional times to develop a thorough description of the setting. Research staff collecting field notes read and discussed each others' typed notes following each observation.

The Instructional Environment Scale–II (TIES-II). The Classroom Observation Protocol of TIES-II (Ysseldyke & Christenson, 1993) was completed after each fall and spring classroom observation. This scale addressed 12 components of effective instruction: (a) Instructional Match, (b) Classroom Environment, (c) Teacher Expectations, (d) Cognitive Emphasis, (e) Motivational Strategies, (f) Relevant Practice, (g) Academic Engaged Time, (h) Informed Feedback, (i) Adaptive Instruction, (j) Progress Evaluation, (k) Instructional Presentation, and (l) Student Understanding. The full TIES-II system includes classroom observation and interviews of teacher, parents, and student, with the intent of problem solving for an individual student. We modified the procedure, using only the classroom observation section, to assign a classroom environment score to each classroom. After observing in the classroom, we completed the TIES-II classroom protocol, scoring each component on a Likert scale from *not at all* (0 points) to *very much* (4 points) characteristic of effective instruction. For each observation, 48 points were possible. Observations of target students were conducted by two researchers until reliability was achieved. We considered our observations reliable if 90% agreement was achieved across three observations. Ninety percent agreement was defined as +/-4 points on the 48-point scale. Reliability was consistently obtained on the first three consecutive observations. After achieving reliability, researchers observed in classrooms alone, typing field notes after each observation. Each observer read the other observers' field notes and completed an independent TIES-II. The observers discussed their ratings after every observation to arrive at agreement on each component and on the total score. The sum of fall and spring ratings was used in analysis.

Fall teacher interviews. During these 1-hour interviews, a semistructured protocol was followed to (a) summarize the teachers' general philosophies towards planning, reading instruction, and assessment; (b) understand teachers' beliefs on learners at-risk for reading difficulties and reading instruction for these students; and (c) corroborate the objectives of the observed lesson. All interviews were tape recorded and transcribed.

Collaborative consultation meeting notes. At each meeting, we followed a protocol to learn about the student identified as DD. Notes were taken on a laptop computer. The protocol included (a) information on the student's background; (b) the students' strengths and needs; (c) the assessment summaries from CBM, phonological awareness tasks, and other pertinent assessment informa-

tion; (d) a list of possible interventions for implementation in the general education classroom; and (e) a description of the chosen intervention, including steps to be followed, materials needed, the frequency of implementation, size of the instructional group, and motivational strategies. A follow-up meeting occurred after 8 weeks of the intervention to assess student progress and to determine the next course of action.

Intervention fieldnotes. Fieldnotes were developed during announced (support check) and unannounced (fidelity check) observations at times when the instructional plan was scheduled to occur. Midway through the intervention plan, a 4-week meeting occurred between the teacher and a researcher to discuss the students' growth as indicated by CBM progress-monitoring data.

Teacher background. Teachers reported background information including the total years of teaching, ethnicity, and earned educational degrees. Total years teaching was used in the quantitative analysis.

Procedures

Prior to the study's inception, each procedure involving human subjects was reviewed and approved by the Institutional Review Board at the University of Maryland.

Child measures. Demographic data and intelligence scores were collected in Year 1. Phonological processing measures were collected in the fall of each year. The WJ-R was administered in late spring of each year. Following the fall screening on CBM measures and receiving parent permission, administration of weekly CBM oral reading fluency measures began, typically in late November. The exception was Year 1 for first grade children who began weekly CBM assessments in January because the task would be too difficult earlier in the year. Teachers completed the SSRS between January and February of each year. Dual discrepancy status was assessed three times a year with the exception of Year 1, in which second grade children were assessed four times. As children were identified as DD, collaborative consultation meetings were scheduled. Three graduate students collected the child measures.

Context measures. One of the authors and a doctoral level graduate assistant observed reading instruction in each classroom during the fall and spring of each year. In addition to fieldnotes, they completed TIES-II after the observation to help analyze and quantify the effectiveness of instruction. They conducted yearly interviews with each teacher following the fall observation. At the end of the year, the teachers completed the background form documenting their formal education and years of experience, and the special educator and/or counselor completed the school attention form.

Data Analysis

Quantitative analysis. Repeated measures analyses of variance were conducted with Group (NDD, IDD, FDD) as the between-subjects factor and Year (1, 2, 3) as the within-subjects factor. These analyses were conducted on the reading, reading-related, behavioral, and instructional context measures using the Proc GLM procedure in SAS (SAS Institute Inc., 1999). Univariate F tests are reported for Group effects and multivariate F tests based on Wilks's lambda criterion are reported for Year and Group x Year interactions. The number of statistical tests conducted raised concerns about committing a Type I error and the small N raised concerns about committing a Type II error. Because most of the measures were grouped into one of four domains, we adjusted the p value within domains to .01 (.05/3 measures per domain) as a compromise between the two error types. We were most interested in Group and Group x Year interactions so multiple comparison follow-up tests using the Tukey-Kramer adjustment ($p < .05$) were conducted if these effects were significant. For completeness we report the main effect for Year but did not conduct follow-up tests because it was expected that most scores would increase over time. Effect sizes were calculated for each variable using Cohen's (1988) d statistic with a pooled standard deviation as recommended by Dunlap, Cortina, Vaslow, and Burke (1996) for repeated measures analysis.

The groups were first compared on age, $F(2, 33) = .77$, intelligence, $F(2, 33) = 2.11$, $p = .14$, and mother's years of education, $F(2, 30) = .58$. The groups were not significantly different on these variables so statistical adjustment was not indicated.

Qualitative analysis. One of the noted challenges of qualitative research is organizing and reducing voluminous data. For this reason, data were first read, reread, and coded. Codes were organized into categories capturing multiple descriptive codes and were used to develop interpretive themes. While developing codes and categories, we wrote and discussed descriptions of the 8 students selected from the IDD and FDD groups. In developing student profiles, we gathered data across the multiple sources where the student may have been discussed or observed (e.g., field notes, interviews, assessment results, collaborative consultation meetings). For each student, we developed two chronological summaries. First, we developed a narrative summary for each student, consisting of detailed memos for data from Years 1 through 3. Next, based on these summaries, we developed charts that depicted critical elements (e.g., student background, achievement, learner characteristics), general education interventions, and classroom context (e.g., teacher background, beliefs, classroom management).

Through the integration of longitudinal quantitative (assessment scores) and qualitative (field notes, interviews) data, we were able to develop a profile for each student. This mixed methodology offered the opportunity to triangulate findings across data sources (Li & Zercher, 1997). For example, we were able to describe a student's performance through curriculum-based measures; through classroom observations, as documented in fieldnotes; and through teacher observations, as documented in the performance ratings (SSRS), teacher interviews, and collaborative meeting notes.

Through discussion, researchers proposed and refuted hypotheses to explore the complex relationship between the instructional environment and child characteristics. Analysis moved back and forth between description, discussion, and interpretation (Merriam, 1988). During the analysis process, we developed additional displays to facilitate identifying similarities and differences across case students' instructional milieu and child characteristics (Bernard, 1988; Miles & Huberman, 1994). Initial themes representing individual students gave way to higher level themes that cut across students.

Results

Results from the analyses of group differences are presented first followed by the qualitative findings.

Group Differences

Reading and behavioral measures. Descriptive statistics for each Group by Year are presented in Table 2 for reading, phonological processing, teacher ratings of classroom behavior, and school attention[2]. To overview the findings, the repeated measures ANOVAs yielded a significant main effect for Group for all measures except Passage Comprehension, Total Phonological Awareness, and RAN-Objects. Word Reading Efficiency was the only variable that yielded a significant Group x Year interaction. In all cases involving a significant Group effect, the FDD group performed significantly worse than the NDD group. Table 3 presents the effect sizes (ES) for each variable.

Specifically, for the subtests of the WJ-R, there were significant main effects for Group and Year, respectively, for Letter Word Identification, $F(2, 32) = 7.82$, $p < .0017$; $F(2, 31) = 126.30$, $p < .0001$ and Word Attack, $F(2, 32) = 5.64$, $p < .008$; $F(2, 31) = 52.87$, $p < .0001$. The Group effect for Passage Comprehension was not significant, $F(2, 32) = 4.99$, $p < .013$, but the Year effect was significant, $F(2, 31) = 50.97$, $p < .0001$. Follow-up tests for Letter Word Identification showed that FDD had poorer skills than IDD and NDD. For Word Attack, FDD < NDD. The ES between FDD and each of the other two groups are considered large differences (i.e., > .80).

The analyses of the phonological processing battery included Total Phonological Awareness, RAN-objects, and Word Reading Efficiency. For Total Phonological Awareness, the main effect for Group was not significant,

Table 2
Descriptive Statistics for Reading, Phonological Processing, Classroom Behavior, and School Attention

	Year 1					
	NDD		IDD		FDD	
	M	*SD*	*M*	*SD*	*M*	*SD*
Letter Word Identification	109.3	9.3	97.7	9.4	89.7	8.6
Word Attack	100.0	9.2	93.8	9.1	87.1	4.2
Passage Comprehension	108.8	10.5	96.9	10.9	88.0	5.6
Total Phonological Awareness	19.8	10.9	16.1	9.6	15.6	11.3
RAN-Objects	41.3	11.2	48.2	13.8	53.4	11.4
WRE	12.1	11.0	12.0	10.5	6.2	7.4
Academic Competence	26.2	5.5	21.9	5.3	16.7	5.7
Social Skills	20.6	2.3	18.2	2.8	19.1	3.6
Problem Behaviors	8.1	5.0	11.1	6.0	13.7	5.5
School Attention	.3	.7	.8	.7	1.9	1.2
	Year 2					
	NDD		IDD		FDD	
	M	*SD*	*M*	*SD*	*M*	*SD*
Letter Word Identification	113.3	11.0	98.7	7.9	87.8	8.1
Word Attack	103.2	13.6	96.8	8.8	85.5	6.1
Passage Comprehension	113.7	9.6	104.6	9.8	95.0	12.9
Total Phonological Awareness	31.8	8.1	28.7	7.9	22.3	9.5
RAN-Objects	34.7	7.8	39.8	9.5	45.4	11.7
WRE	42.9	14.8	32.9	15.2	14.6	9.4
Academic Competence	31.0	4.7	25.2	5.8	18.7	3.6
Social Skills	23.8	4.5	20.9	5.2	15.0	5.8
Problem Behaviors	5.0	5.6	8.5	5.1	17.3	9.1
School Attention	0	0	.2	.8	1.14	1.1

(Table 2 continues)

(Table 2 continued)

	Year 3					
	NDD		IDD		FDD	
	M	*SD*	*M*	*SD*	*M*	*SD*
Letter Word Identification	108.9	10.1	100.8	8.0	85.4	7.5
Word Attack	103.4	16.3	93.7	8.0	82.0	11.3
Passage Comprehension	112.7	9.8	104.5	11.1	89.6	13.9
Total Phonological Awareness	33.5	8.2	33.1	7.0	28.0	7.6
RAN-Objects	30.4	5.6	34.8	6.5	36.3	5.3
WRE	59.1	9.6	49.3	10.6	26.9	13.3
Academic Competence	30.4	5.3	24.4	5.3	18.0	3.7
Social Skills	21.3	5.2	20.1	5.2	16.0	3.8
Problem Behaviors	6.3	4.7	8.6	4.6	12.4	3.5
School Attention	0	0	.2	.4	2.0	3.3

Note. NDD = At-risk, not dually discrepant; IDD = At-risk, infrequently dually discrepant; FDD = At-risk, frequently dually discrepant, WRE = Word Reading Efficiency.

F (2, 33) = 1.55, $p < .22$, but, as expected, the Year effect was significant, F (2, 32) = 81.36, $p < .0001$. Similar results were obtained for RAN-Objects with a nonsignificant effect for Group, F (2, 32) = 2.90, $p < .07$ and a significant effect for Year, F (2, 31) = 45.23, $p < .0001$. Despite the lack of statistical significance for Group, the ES for these two variables were interpretable and ranged from small (-.25) to large (-.90) differences.

Word Reading Efficiency yielded a significant main effect for Group, F (2, 33) = 9.35, $p < .0006$, Year, F (2, 32) = 311.78, $p < .0001$, and a significant Group x Year interaction, F (4, 64) = 9.73, $p < .0001$. Multiple comparisons within year were conducted to explain the interaction. There were no statistically significant group differences for Year 1. Group differences were obtained for Year 2 between the FDD group and both IDD and NDD (FDD < IDD, NDD). For Year 3, the significant group comparisons were FDD < IDD < NDD. The ES reported in Table 3 are moderate for Year 1 but then become substantial for Years 2 and 3 and are indicative of a fan spread interaction effect in which the groups become more disparate over time.

Teacher ratings of Academic Competence yielded a main effect for Group, F (2, 32) = 19.82, $p < .0001$, and Year, F (2, 31) = 7.59, $p < .003$. Follow-up tests indicated all group comparisons were statistically significant with FDD < IDD < NDD. Ratings of Social Skills yielded a significant main effect for Group, F (2, 32) = 8.01, $p < .0015$, but not Year, F (2, 31) = .17. FDD had significantly poorer social skills compared to NDD. Ratings of Problem Behaviors yielded a significant main effect for Group, F (2, 32) = 13.00, $p < .0001$, but not Year, F (2, 31) = 2.71, $p < .09$, with all follow-up comparisons significant, FDD > IDD > NDD. The ES in Table 3 suggest moderate to large differences.

School Attention captured extra-classroom concerns about the children. There was a significant main effect for Group, F (2, 33) =

Table 3
Effect Sizes and Statistical Significance for Groups on Reading, Phonological, Behavioral, and Contextual Variables

		Group Comparisons		
		NDD-FDD	IDD-FDD	NDD-IDD
Letter Word Identification		1.37*	.87*	.49
Word Attack		1.32*	1.05	.38
Passage Comprehension		1.08	.68	.39
Total Phonological Awareness		.62	.38	.25
RAN-Objects[a]		-.90	-.36	-.51
Word Reading Efficiency[b]	Y1	.64	.65	.00
	Y2	2.34*	1.49*	.67
	Y3	2.89*	1.89*	1.05*
Academic Competence		2.33*	1.23*	1.01*
Social Skills		1.20*	.65	.54
Problem Behavior[a]		-1.41*	-.86*	-.60*
School Attention[a]		-1.56*	-1.20*	-.50
TIES-II		.17	.13	.04
Teacher years of experience		-.21	-.04	-.15
Classroom reading slope		-.42	-.39	-.04

Note. NDD = At-risk, not dually discrepant; IDD = At-risk, infrequently dually discrepant; FDD = At-risk, frequently dually discrepant; TIES-II = The Instructional Environment System-II.
*$p < .05$.
[a]These variables are reversed scored so negative effect sizes favor the first group listed at the top of the column. [b]Because of the significant Group x Year interaction, effect sizes are reported by year (Y1, Y2, Y3).

9.95, $p < .0001$, and for Year, $F(2, 32) = 5.97$, $p < .007$. Follow-up tests indicated significant differences between the FDD group and each of the other two groups, FDD > IDD, NDD. Both ES for FDD were greater than 1.0.

Instructional context. Ratings of the instructional environment (TIES-II), years of teaching experience, and reading progress of peers were used to examine differences in instructional context. Table 4 presents the descriptive statistics for these variables and Table 3 presents the ES. The repeated measures ANOVA for TIES-II was not significant for Group, $F(2, 33) = .19$, but the Year effect was significant, $F(2, 32) = 6.70$, $p < .004$. The Group x Year interaction was not significant, $F(4, 64) = .22$. The years of teaching experience analysis yielded a nonsignificant effect for Group, $F(2, 33) = .36$, but a significant main effect for Year, $F(2, 32) = 7.45$, $p < .0022$. The interaction was not significant, $F(4, 64) = 1.41$, $p < .25$. The analysis of classroom slope yielded no significant main effects for Group, $F(2, 33) = 2.43$, $p < .11$, Year, $F(2, 32) = 3.09$, $p < .06$, or the interaction, $F(4, 64) = 1.39$, $p < .25$. Thus, there were no quantitative group differences or interactions for instructional context as measured in this study. Seven of the

Table 4
Descriptive Statistics for Contextual Variables

	Year 1						Year 2						Year 3					
	NDD		IDD		FDD		NDD		IDD		FDD		NDD		IDD		FDD	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
TIES-II	82.8	10.5	82.0	10.2	82.0	11.2	85.9	4.5	85.8	4.9	85.4	4.6	82.3	5.7	82.2	5.0	79.6	6.9
Teacher years of experience	13.4	10.3	14.9	9.4	12.0	6.4	6.0	6.4	9.1	8.7	14.0	10.9	5.4	7.6	5.1	8.3	4.1	4.5
Classroom reading slope	1.16	.21	1.15	.23	1.10	.25	.92	.22	.95	.13	1.17	.21	.95	.28	.95	.29	1.10	.38

Note. NDD = At-risk, not dually discrepant; IDD = At-risk, infrequently dually discrepant; FDD = At-risk, frequently dually discrepant; TIES-II = The Instructional Environment System-II.

nine ES for this group of variables were less than .20, which is not considered meaningful.

Qualitative Differences

To better understand the quantitative findings that students in the FDD group have weaker reading, reading-related, and behavioral skills, we explored how individual student characteristics and classroom environments explained or contributed to students' reading progress. The quantitative analyses revealed individual differences important to student progress, but indicated no differences across instructional contexts. Yet, our extended involvement in the school allowed us a closer look at the setting and provided qualitative data that afforded a more in-depth, interactive examination. Therefore, our qualitative analysis was initially intended to spotlight classroom contextual features that we hypothesized differentiated student performance for 8 case students (see Table 5).

First, we examined individual student characteristics and classroom environmental features in isolation. We described each student's "persona," defined by Webster's Dictionary as the "outer personality or façade presented to others by an individual." To determine persona, we considered a student's background, academic skill progress and level, home support of school, home-school connection, and in-school disposition. Similar to the quantitative analyses of teacher ratings of classroom behavior, we found that IDD students were more engaged in their own learning than FDD students. Teachers and research staff viewed IDD students as more willing to take risks, more confident in class, and more motivated in their academics than FDD students. Yet, persona did not appear to be the sole distinguishing feature between the two groups.

Over 3 years, we examined several classroom features, keeping the individual student as the focus of analysis. For each teacher, we analyzed management style, interactions with students, and reading instructional methods, including grouping practices and individualization of instruction. Finally, we considered the teacher's overall commitment to the research project and willingness to modify instruction for students identified as DD. Across 3 years, 3 of the 4 students in both the FDD and IDD groups were placed in moderate to strong instructional environments. Therefore, by examining persona and environment in isolation, we did not arrive at a satisfactory explanation of possible qualitative differences between the FDD and IDD children.

Following a long line of theory and research, we rediscovered that environments are socially constructed (Bronfenbrenner, 1979; Gallimore, Goldenberg, & Weisner, 1993; Vygotsky, 1978), and found that interactions between students' personae and their environments differentiated the FDD and IDD groups better than either element in isolation. We examined interactions between students and their classroom environments with the idea of better understanding how students accessed their instructional milieu.

Access

We found a student's ability to enter into or make use of the reading curriculum—our definition of access—was the central theme distinguishing the FDD and IDD groups. Persona and environment worked together to determine how well students accessed learning. To some extent, we viewed the interaction as summative, so that Persona + Environment = Access. When either persona or environment was less than optimal, the other addend had to compensate to ensure that a student could access learning. Overall, if P = 1/2 and E = 1/2, IDD students came closer to the equation summing to 1 more frequently than did FDD students. Optimal access, and therefore greater learning, occurred as the sum approached 1, with both persona and the instructional environment supporting students' access to learning.

Evaluating the environment. The concept of access was more complex than the equation implies, however, as the addends were not mutually exclusive. In order to evaluate a student's ability to access the instructional environment, we found it necessary to evaluate the context in light of the match between the student's persona and the classroom environment.

Table 5
Description of Case Students

Student	Status	Grade[a]	Race	Gender	Language[b]	Reading	IQ	Access
Janis	FDD	1	HS	F	Spanish	93	71.1	Low-Moderate
Victor	FDD	1	C	M	English	88	107	Low-Moderate
Olanda	FDD	2	AA	F	English	82	85.6	Low-Moderate
David	FDD	2	HS	M	unspecified	88	88.1	Low
Keri	IDD	1	HS	F	Spanish	90	69.6	Moderate
Thomas	IDD	1	A	M	unspecified	106	101.7	Moderate
Johanna	IDD	2	HS	F	Spanish	105	92	Moderate-High
Kyle	IDD	2	C	M	English	102	113.6	High

Note. For Status, FDD = frequently dually discrepant, IDD = infrequently dually discrepant; For Race, A = Asian, AA = African American, C = Caucasian, HS = Hispanic; Reading = Woodcock-Johnson Psycho-Educational Battery—Revised, Basic Reading standard scores Year 3; IQ = Wechsler Intelligence Scale for Children—Revised Full Scale IQ.
[a]Grade at the beginning of the study. [b]Language spoken at home as reported by teachers.

Sometimes persona overrode what we perceived as an excellent instructional context. For example, in the second and third grades, Olanda—identified FDD—was in classrooms that we felt were outstanding because of strong, veteran teachers with excellent management and instructional skills. However, we rated Olanda's environments as less than optimal because they failed to motivate her. Olanda's second- and third-grade teachers perceived Olanda as being more interested in gaining social acceptance than in academics. In fourth grade, however, when Olanda was in the classroom of another excellent teacher, who was less experienced, but young and energetic, and who offered social opportunities during recess and after school (e.g., help with talent show routines), Olanda was not identified DD. The fourth-grade classroom appeared to provide a better match for her personality.

Further, a student's persona could improve the rating of an instructional environment. For example, although we rated one third-grade classroom positively overall, we rated it particularly high for Keri—identified IDD. The teacher recognized Keri's positive attitude and effort and, in turn, enriched Keri's environment by referring her for added reading support. The following sections describe the four possible persona/environment configurations observed.

Strong persona, strong environment. As previously noted, optimal access occurred when both the student's persona and the environment worked in concert. Kyle—identified IDD—was in second grade in the study's first year. Across 3 years, teachers described Kyle as impulsive but also as a risk taker. In second grade, he completed independent tasks quickly but incompletely, and spent much of his time out of his seat with reminders of where to be and what to do. He continued displaying this impulsivity in third and fourth grades. Yet, Kyle appeared interested in classroom activities across 3 years. In small groups, he demonstrated improved focus and answered questions accurately. On several occasions he responded correctly to questions although he appeared off-task. We observed steady improvement in Kyle's reading performance across the three years of the study.

Interactions between Kyle's persona and classroom contexts may help explain this improvement. In the second grade, Kyle had a novice teacher. She was willing but unable to individualize instruction, possibly due to her lack of experience in classroom management and pedagogical knowledge. Yet, Kyle's parents offered a presence (his father volunteered in the classroom regularly) and support that enabled the school to identify his strengths and weaknesses and to plan for slightly specialized services in the regular education setting. In Grades 3 and 4, Kyle's teachers were adept at working with him. His third-grade teacher set firm, but fair, limits on his classroom behavior and focused on learning. For example, she shared with Kyle graphs of his reading progress developed by the research staff and tapes of his oral reading to provide feedback and motivation to improve. He encountered a similar environment in fourth grade with a teacher who espoused high academic expectations for her students and insisted on their best work while demonstrating skill in meeting students' individual needs. Observations of reading instruction and small group interventions in Kyle's third- and fourth-grade classrooms revealed an excellent match between Kyle and his instructional environments, allowing attention to his strengths and compensation for his impulsivity.

Strong persona, weak environment. With a weak instructional environment, student persona had to compensate in order for the student to access learning. Johanna (IDD) was in Grade 2 at the start of the study. Teachers' perceptions of Johanna varied across the 3 years. In second grade, Johanna's teacher described her as "worldly wise" and "proud," unwilling to "admit when she doesn't know anything." Further, the teacher explained that Johanna did not seem to care about pleasing others and did not demonstrate that school had any value.

Our observations of the second-grade classroom context were negative. Although the teacher made efforts to incorporate elements of good instruction, the classroom in general had little structure. Lesson objectives and expectations were unclear, and poor behavior management contributed to the teacher using a negative tone. As a collaborative team, we

encountered numerous roadblocks working with the teacher. She resisted working with smaller instructional groups, stating that this "took away from the whole class."

Despite this difficult year, Johanna made reading progress that pulled her out of DD status before the end of the year. She finished the year with a slope of 1.22 and reading on average 60 words per minute (wpm), compared to second-grade median slope and wpm values of 1.09 and 66.1, respectively. In the following 2 years of the study, the second-grade teacher's description of Johanna being "proud" was turned into an asset. In third grade, we observed Johanna as "comfortable and confident enough to ask questions...as well as express her own ideas." During third and fourth grades, Johanna steadily gained academic independence, confidence, and skill. By the end of the fourth grade, her oral reading fluency measured 102 wpm with a slope of 1.45. These scores compared favorably with the fourth-grade median score of 133.1 wpm and median slope of 1.23.

Weak persona, strong environment. With a strong environment and a weak student persona, we observed the teacher/environment seeking out the student. If the student persona did not respond, however, access (and subsequently student achievement) would not follow. We discuss Janis and Victor who fit this pattern.

Janis was in first grade at the beginning of the study. In all 3 years, she was identified as DD during some part of the school year. Teachers described her as "nice," "really quiet," and a hard worker. However, overriding these positive characteristics, teachers described Janis as "slow." Despite being cooperative, Janis was consistently the last student to join her reading group. In second grade, when Janis received increased teacher support and appeared most motivated, teachers reported that she had a good year, but was "still slow." Throughout the 3 years of the study, Janis received individualized interventions collaboratively developed by her teachers and the research team. Interventions involved phonological awareness, the alphabetic principle, sight word fluency, and reading connected text. Janis demonstrated growth over the 3 years. However, her growth was not enough to change her DD status. Her year-end oral reading fluency slopes ranged from .47 in Year 1 to .59 in Year 3.

Like Janis, Victor fell into the weak persona, strong environment category. Victor began the study in first grade. From the start, Victor was identified DD; he performed significantly worse than his classmates throughout the study. Victor's first-grade teacher described him as hypoactive, off task, and requiring constant monitoring to complete and turn in assignments. She stated that he "seems brighter than his performance indicates." In April, Victor's first-grade teacher referred him for special education evaluation. Victor was identified with LD and Attention Deficit Disorder (ADD). Across the years, Victor received small group or individual help from many sources, including his teachers, instructional assistants, and special education teachers. This instruction included: small-group emphases in first and second grades on phonemic awareness and phonics skills; individual instruction on a modified Glass Analysis procedure (structural word analysis), followed by oral reading practice, summarizing, and follow-up written comprehension questions; and pull-out special education services three to four times a week by the end of first grade and during second grade. Our measures indicated that Victor demonstrated reading progress; however, it was consistently slower and at a lower level than his classmates. For example, his oral reading fluency wpm for Grades 1 through 3, respectively, were 5, 13, and 34 compared to the median values for his peers of 38.3, 66.1, and 109.

Weak persona, weak environment. Most frustrating to us were cases where students' weaknesses were exacerbated by weak classrooms. David (FDD) was in second grade at the beginning of the study. We evaluated his second- and third-grade classrooms as weak for students experiencing reading difficulty. In second grade, David was in the same class as Kyle. His teacher was a novice, willing but unable to plan and carry out individualized instruction for students falling behind the curriculum. Unlike Kyle, whose parents offered a presence and support

that enabled Kyle to succeed, David's mother was involved for disciplinary actions but was not consulted, to our knowledge, to support David's reading. In third grade, David's teacher was veteran. The principal described this teacher as strong, but lacking in her ability or desire to work with students with diverse needs. We corroborated this view from our observations and an interview. We found it difficult to secure her participation at the collaborative consultation meetings despite having a substitute teacher assigned to cover her class. When an intervention was developed, the intervention happened rarely and had little chance of impacting David's performance. In contrast, David was in the same fourth-grade classroom as Kyle and Olanda, with a teacher who espoused high academic expectations for all of her students and worked to meet individual needs. For example, she gave David individualized attention soon after the start of an independent writing activity. She had him think of a transition sentence and motivated him to remember minor details to support his sentence. When writing time ended, David appeared anxious because he was not finished. His teacher assured him that she would find time later to work with him. David's oral reading fluency improved markedly from third grade to fourth grade. Similar to Olanda, the connection between his persona and environment appeared to accelerate growth. Comparing third grade to fourth grade, he progressed from reading 33 to 86 wpm in more difficult text, with a growth in slope from .51 to 1.33 words per week.

Summative impact of access. The previous sections illustrated how combinations of persona and instructional environment affected access for individual students. After examining these configurations, we developed a summary access rating for each student. Two of the authors independently evaluated each student's access, rating each student on a continuum from low to high access. The two authors' evaluations were in agreement, and the third author confirmed their judgments. We found that IDD students had higher access levels than FDD students (see Table 5, last column).

Discussion

We examined the validity of a response-to-instruction model to identify children who may benefit from special education services. Using a mixed methodology design, we studied both child characteristics and classroom contextual features to examine differences between children exhibiting different degrees of responsiveness to general education reading interventions. We hypothesized that children who demonstrated persistent nonresponsivness (FDD) would also demonstrate more severe weaknesses on reading and behavioral skills and would receive more attention from sources beyond the classroom. We also examined whether the most unresponsive children experienced less ideal classroom settings.

Validity From the Perspective of Individual Differences and Context

The interpretation of the individual difference analyses is clear. Persistent nonresponsiveness to general education instruction as defined by FDD status correlated with deficits on reading skills and teacher ratings of behavior. Further, FDD children were more likely to receive additional concern at the school level. Together, these findings indicate that the response-to-instruction model possesses both criterion and social validity.

The FDD children scored lower on all measures, compared to children initially deemed at-risk but never identified as dually discrepant (NDD). These differences were substantial as reflected by a median ES of 1.34. Further, FDD children differed from IDD children on Letter Word Identification, Word Reading Efficiency, Academic Competence, Problem Behaviors, and Social Skills. In fact, on single word reading fluency, the disparity between the FDD children and the other two groups became larger over time. The median ES between FDD and IDD was .865. The differences between FDD and each of the other groups suggest that the focus on both dual discrepancies as measured by CBM and responsiveness to general education instruction is a sensitive and valid indicator of child status.

Our supposition that FDD children may be candidates for special education services is supported by CBM slope estimates provided by Deno, Fuchs, Marston, and Shinn (2001). They estimated that oral reading fluency slopes for children in special education were .71 in first and second grades and .58 in third and fourth grades. The FDD children's mean slopes were .28 (first/second grades) and .67 (third/fourth grades), which approximate these figures. Although the FDD children would, by definition, have the lowest slopes of the children in this study, the fact that their slopes reflect estimates for diagnosed samples of children with LD provides further evidence in support of criterion-related validity.

A pervasive issue in identifying LD in schools is the incredible growth in the number of children identified. Over a 10-year period, the number of children identified increased by almost 40%, representing 4.45% of the school-age population, while other high-incidence categories remained stable (ERIC Clearinghouse on Disabilities and Gifted Education, 2001). Thus, an important issue in considering an alternative identification method is the number of children who would be identified. Because we derived our sample from a population screen, we can estimate directly the percentage of children who may be expected to fall into our groups. The FDD group represented 2.8%, IDD = 6.8%, and NDD = 4.8% of first- and second-grade children. These are almost certainly underestimates for children in elementary school because the at-risk group, which constituted the longitudinal sample, was identified in Year 1. Likely, more children would have been identified as DD in ensuing years had our design and resources allowed for this possibility.

The analyses of classroom context did not support the hypothesis that FDD children received poorer quality instruction, had peers who demonstrated poorer oral reading fluency growth, or had less-experienced teachers. In fact, the ES estimates suggest that the FDD group had a slight advantage in teacher experience and classroom reading slope. However, these indicators may be too global to index child-environment interactions known to vary within classroom (Keogh & Speece, 1996). In any event, the differences observed between the groups cannot be attributed to differential classroom environments as measured in this study, lending further credence to the validity of the classifications.

Access to Learning: Beyond Individual Differences and Context

The qualitative findings provided an opportunity to examine more closely potential contextual influences. We began the qualitative analysis with an exclusive focus on instructional features, but discovered that more than context was operating in the distinction between FDD and IDD children. We concluded that the 4 FDD children experienced poorer access to learning, defined as the reciprocal interactions between the child's persona and instructional environment. This is not a new concept in the developmental literature on learning and competence, but it is often not a focus of research on early reading (Lipson & Wixson, 1986). The developmental literature has emphasized the importance of assessing domains beyond cognitive competence to include adjustment and motivation (Scarr, 1981), the child's engagement with school aims and values (Maughan, 1988), and goodness of fit between child temperament and the context (Rothbart & Jones, 1999). Our conception of access is similar, but extends the reasoning to include the adjustment of the context (e.g., teacher, instruction) in addition to the child's motivation and skill in adjusting to the context encountered. For example, Johanna and David experienced strong improvement in fourth grade with a teacher who was able to spark their interest in literacy. It is not possible to pinpoint the interactional dynamics that fueled increased access for these children, given the limits of the data set. Also, we cannot rule out increased maturity and development on the children's part as an explanation for their growth. Nonetheless, our findings suggest that closer attention to the child-context fit may net increased understanding of child success and failure with implications for research and instruction.

A point of congruence between the qualitative and quantitative findings is access to learning and teacher ratings of behavior. The three groups of children were distinguishable on the behavioral ratings. The behavior ratings may reflect access from the teacher's perspective, incorporating both her view of the child in context and her ability to adjust to the child's unique needs. Interpreting ratings of behavior through the lens of "access to learning" assists in explaining the strong and long-standing relationship between classroom behavior and academic achievement (e.g., Kavale & Forness, 1995; McKinney, Mason, Perkerson, & Clifford, 1975; Torgesen et al., 1999).

Level and rate of progress, as indexed by dual discrepancies, may provide an even more direct measure of access than behavior ratings. Dual discrepancies are composed of both child abilities and teacher skill and thus transcend the dichotomy of individual differences and contextual effects. A child who is dually discrepant from peers may lack skills, experience a less-than-optimum instructional environment or, most likely, both. As demonstrated in this study, some children will respond to instructional improvements in general education (IDD children) and others will not (FDD children). We wish to emphasize that the problem does not belong entirely to general education. For some children a good general education environment is not enough. Our descriptions of Janis and Victor illustrate this point. Despite the well-informed and determined efforts of many teachers, these children could not keep pace with their peers.

Conclusions and Implications

A modified response-to-instruction paradigm originally operationalized by Fuchs and Fuchs (1998) identified a valid group of children who are in need of special education services. Persistently nonresponsive children differed from the other at-risk groups on a variety of theoretically and instructionally relevant measures, and could not access learning within the general education environment despite specially designed interventions. These conclusions need to be viewed within the context of the study's limitations. The generalizability of the findings is limited by the facts that we studied a single school, focused only on the academic domain of reading, and included only children from first through fourth grade. Further work with different populations (e.g., rural, inner city, younger and older children) and more academic domains (e.g., math, writing) is necessary to examine further the validity of the model. A strength of the study was the use of a population screen to identify an at-risk group from which DD children were identified. This design feature avoided sample selection bias present in school-identified samples (Shaywitz, Shaywitz, Fletcher, & Escobar, 1990; Speece & Case, 2001) and is essential to the examination of alternative identification approaches.

A second issue is that we used a 3-year time period to identify responsiveness to instruction. Although the longitudinal design is a strength from many perspectives, the question remains whether response to instruction as an approach to identifying LD can be implemented within a single year. Further research on implementing the model within an academic year, provided that progress monitoring can begin in September, and paralleling the comparisons made in the current study will be critical to judgments of validity.

An important aspect of this study is that general education staff within their own classrooms conducted the interventions. In contrast, other studies focusing on early reading have implemented intensive tutoring by researcher-trained instructors (e.g., Torgesen et al., 1999; Vellutino et al., 1996) following the identification of an at-risk sample. These efforts, which are representative of what should go on in special education, are targeted more at prevention than identification. Advantages of the model we implemented are the potential to empower general education teachers with knowledge of research-based instruction and the potential to improve instruction and, hence, performance of a significant number of children. The IDD children in our study represented almost half of our at-risk sample and constituted a "gray area" as they went in and out of risk over the course of this study. The response-to-instruction model provides a

method of monitoring these children and their instructional environments so they do not become lost in the system.

The response-to-instruction model improves upon the current prereferral system in several ways. First, all children are monitored and evaluated continuously with respect to peers, eliminating potential referral bias and the problem of false negatives (i.e., children who should receive intervention but are missed by current methods). Further investigations of classification accuracy will be necessary to determine the extent of classification errors with a response-to-instruction model. Second, the measurement model, based on CBM, is well established. Third, intervention implementation is monitored for fidelity (reliability) of treatment. A fourth strength, not examined in this study, is that overall low class performance (i.e., level and slope) compared to other classrooms would indicate that classroom-level intervention is required prior to identifying children as DD (Fuchs & Fuchs, 1998).

The qualitative analysis of "access" assisted understanding of the role of child-environment fit in school achievement. Although this interaction may be difficult to operationalize quantitatively, a starting point may include measurement of temperament, motivation, home support of schooling, and home-school connections as an attempt to capture what we called "persona." Relevant contextual variables could be extended beyond our use of TIES-II (Ysseldyke & Christenson, 1993) to include other ecobehavioral variables as represented in Greenwood's (1996) work.

The LD field is at a critical juncture with respect to its future. Both logic and evidence point to responsiveness-to-instruction as a centerpiece of new ways of conceptualizing LD (Donovan & Cross, 2002; Fuchs & Fuchs, 1998; Speece & Shekitka, 2002). We provided evidence that the model we evaluated can be implemented within a general education framework to identify a group of children who differ from their peers on important reading success indicators and who require more intensive instructional efforts. The validity of other approaches to identification requires examination, as does a response-to-instruction model with a broader sample and attention to other academic domains.

Footnotes

[1] Because fidelity is central to the concept of treatment responsiveness, we tested yearly differences between the two groups who received interventions (IDD, FDD) to determine if either group received more faithful instruction, which might explain different outcomes. Two-tailed t tests within year revealed no significant differences between the groups (all t values $\leq .66$; all p values $\geq .52$). Differences in treatment fidelity do not appear to be a viable explanation for group differences identified in this study.

[2] Examination of Table 2 shows that the spring mean performance of the NDD group (and, to a lesser extent, the IDD group) on the WJ-R reading subtests were at or above the normative means for the measures even though the children met our criterion for at-risk status in the fall (lowest 25% of children). This criterion was used to identify a longitudinal sample likely to experience difficulty and not as an early identification measure. Nonetheless, we examined the fall screening data to determine if the groups differed. For the first-grade children, a one-way ANOVA indicated no group differences on fall LSF, $F(2,18) = 1.38$, $p = .24$. For second-grade children, the ANOVA for fall ORF was significant, $F(2,12) = 4.83$, $p = .03$. Follow-up tests with Tukey's HSD showed that FDD < NDD but no differences involving IDD. All 3 FDD second-grade children scored below the 10th percentile, calculated on the population screened ($N = 127$). Of the 4 NDD second-grade children, 2 scored below the 25th percentile (16th and 14th) and 2 scored above (29th and 33rd). Thus, although the NDD second-grade group was superior to the FDD second-grade group on the fall ORF screening measure, the NDD group was in the bottom third of all second-grade children screened and it was appropriate to consider them at-risk for reading difficulty in the fall.

References

Benes, K. M. (1995). [Review of the Social Skills Rating System.] In J. C. Conoley & J. C. Impara (Eds.). *The twelfth mental measurements yearbook* (pp. 965-967). Lincoln, NE: The Buros Institute of Mental Measurements.

Bernard, H. R. (1988). *Research methods in cultural anthropology*. Newbury Park, CA: Sage Publications.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments in nature and design*. Cambridge, MA: Harvard University Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219-232.

Deno, S. L. (1989). Curriculum-based measurement and alternative special education services: A fundamental and direct relationship. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 1-17). New York: The Guilford Press.

Deno, S. L., Fuchs, L.S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30,* 507-524.

Donovan, S., & Cross, C. (Eds.). (2002). *Minority students in gifted and special education.* Washington, DC: National Academy Press.

Duncan, G. J., & Raudenbush, S. W. (1999). Assessing the effects of context in studies of child and youth development. *Educational Psychologist, 34,* 29-41.

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods, 1,* 170-177.

ERIC Clearinghouse on Disabilities and Gifted Education. (2001, April). *Educating exceptional children: A statistical profile.* Arlington, VA: The Council for Exceptional Children.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stuebing, K. K., Francis, D. J., Fowler, A. E., & Shaywitz, B. A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology, 86,* 6-23.

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-Assisted Learning Strategies: Making classrooms more responsive to academic diversity. *The American Educational Research Journal, 34,* 174-206.

Fuchs, L. S. (1995, May). *Incorporating curriculum-based measurement into the eligibility decision-making process: A focus on treatment validity and student growth.* Paper presented at the Workshop on IQ Testing and Educational Decision Making, National Research Council, National Academy of Science, Washington, DC.

Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research and Practice, 13,* 204-219.

Fuchs, L. S., Hamlett, C. L., & Fuchs, D. (1997). *MBSP Basic Reading manual* (2nd ed.) Macintosh version [Computer software]. Austin TX: PRO-ED.

Furlong, M., & Karno, M. (1995). [Review of the Social Skills Rating System.] In J. C. Conoley & J. C. Impara (Eds.), *The twelfth mental measurements yearbook* (pp. 967-969). Lincoln, NE: The Buros Institute of Mental Measurements.

Gallimore, R., Goldenberg, C. N., & Weisner, T. S. (1993). The social construction and subjective reality of activity settings: Implications for community psychology. *American Journal of Community Psychology, 21,* 537-559.

Gerber, M. (2002). *A meta-research commentary on MacMillian and Siperstein's "Learning disabilities as operationally defined by schools."* In R. Bradley, L. Danielson, & D. P Hallahan (Eds.), Identification of learning disabilities: Research to practice (pp. 341-350). Mahwah, NJ: Erlbaum.

Greenwood, C. R. (1996). Research on the practices and behavior of effective teachers at the Juniper gardens children's project: Implications for the education of diverse learners. In D. L. Speece & B. K. Keogh (Eds.), *Research on classroom ecologies* (pp. 39-79). Mahwah, NJ: Erlbaum.

Gresham, F. M., & Elliott, S. N. (1990). Social Skills Rating System. Circle Pines, MN: American Guidance Service.

Gresham, F. M., & Witt, J. C. (1997). Utility of intelligence tests for treatment planning, classification, and placement decisions: Recent empirical findings and future directions. *School Psychology Quarterly, 12,* 249-267.

Jenkins, J. R., & O'Connor, R. E. (2002). Early identification and intervention for young children with reading/learning disabilities. In R. Bradley, L. Danielson, & D. P Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 99-149). Mahwah, NJ: Erlbaum.

Kavale, K. A., & Forness, S. R. (1995). *The nature of learning disabilities: Critical elements of diagnosis and classification.* Mahwah, NJ: Erlbaum.

Keogh, B. K., & Speece, D. L. (1996). Learning disabilities within the context of schooling. In D. L. Speece & B. K. Keogh (Eds.), *Research on classroom ecologies: Implications for inclusion of children with learning disabilities* (pp. 1–14). Mahwah, NJ: Lawrence Erlbaum.

Li, S., & Zercher, C. (1997, November). *Data analysis strategies and issues in mixed method research.* Paper presented at the annual conference of the American Evaluation Association, San Diego, CA.

Lipson, M. Y., & Wixson, K. K. (1986). Reading disability research: An interactionist perspective. *Review of Educational Research, 56,* 111-136.

MacMillan, D. L., & Siperstein, G. N. (2002). Learning disabilities as operationally defined by schools. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 287-333). Mahwah, NJ: Erlbaum.

Mathes, P. G., Grek, M. L., Howard, J. K., Babyak, A. E., & Allen, S. H. (1999). Peer-assisted learning strategies for first-grade readers: A tool for preventing early reading failure. *Learning Disabilities Research and Practice, 14,* 50-60.

Maughan, B. (1988). School experiences as risk/protective factors. In M. Rutter (Ed.), *Studies of psychosocial risk: The power of longitudinal data* (pp. 200-220). New York: Cambridge.

McKinney, J. D., Mason, J., Perkerson, K., & Clifford, M. (1975). The relationship between classroom behavior and academic achievement. *Journal of Educational Psychology, 67,* 198-203.

Merriam, S. B. (1988). *Case study reseach in education: A qualitative approach.* San Francisco: Jossey Bass.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage Publications.

National Reading Panel. (2000). *Summary report*. Washington, DC: National Institute of Child Health and Human Development.

President's Commission on Excellence in Special Education. (2002). *A new era: Revitalizing special education for children and their families*. Retrieved July 10, 2002, from http://www.ed.gov/inits/commissionsboards/whspecialeducation.

Rothbart, M. K., & Jones, L. B. (1999). Temperament: Developmental perspectives. In R. Gallimore, L. P. Bernheimer, D. L. MacMillan, D. L. Speece, & S. Vaughan (Eds.), *Developmental perspectives on children with high incidence disabilities* (pp. 33-53). Mahwah, NJ: Erlbaum.

Salvia, J., & Ysseldye, J. E. (1995). *Assessment* (6th ed.). Boston: Houghton Mifflin.

SAS Institute Inc. (1999). *SAS/STAT user's guide, Version 8*. Carey, NC: Author.

Sattler, J. M. (1988). *Assessment of children (3rd ed.)*. San Diego, CA: Author.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75-107). Timonium, MD: York Press.

Scarr, S. (1981). Testing for children: Assessment and the many determinants of intellectual competence. *American Psychologist, 36*, 1159-1166.

Schatschneider, C., Francis, D. J., Foorman, B. R., Fletcher, J. M., & Mehta, P. (1999). The dimensionality of phonological awareness: An application of item response theory. *Journal of Educational Psychology, 91*, 439-449.

Share, D. L., McGee, R., & Silva, P. (1989). IQ and reading progress: A test of the capacity notion of IQ. *Journal of American Academy of Child and Adolescent Psychiatry, 28*, 97-100.

Shaywitz, S. E., Shaywitz, B. A., Fletcher, J. M., & Escobar, M. D. (1990). Prevalence of reading disability in boys and girls. *Journal of the American Medical Association, 264*, 998-1002.

Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children*. New York: The Guilford Press.

Siegel, L. S. (1988). Evidence that IQ scores are irrelevant to the definition and analysis of reading disability. *Canadian Journal of Psychology, 42*, 201-215.

Snow, C., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

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.

Speece, D. L., & Shekitka, L. (2002). How should reading disabilities be operationalized? A survey of experts. *Learning Disabilities Research & Practice, 17*, 118-123.

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.

Steubing, K. K., Fletcher, J. M., LeDoux, J. M., Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2002). Validity of IQ-Discrepancy classifications of reading disabilities: A meta-analysis. *American Educational Research Journal, 39*, 469-518.

Torgesen, J. K., & Burgess, S. R. (1998). Consistency of reading-related phonological processes throughout early childhood: Evidence from longitudinal-correlational and instructional studies. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 161-188). Mahwah, NJ: Erlbaum.

Torgesen, J. K., & Wagner R. K. (1997). Comprehensive Test of Phonological Processes (norming version). Austin, TX: PRO-ED.

Torgesen, J. K., & Wagner, R. K. (1998). Alternative diagnostic approaches for specific developmental reading disabilities. *Learning Disabilities Research & Practice, 13*, 220-232.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E. R., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology, 91*, 579-593.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Chen, R., Pratt, A., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88*, 601-638.

Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wagner, R. K., Torgesen, J. K., Laughon, N. P., Simmons, K., & Rashotte, C. A. (1993). Development of young readers' phonological processing abilities. *Journal of Educational Psychology, 85*, 83-103.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology, 30*, 73-87.

Wechsler, D. (1974). Wechsler Intelligence Scale for Children–Revised. New York: Psychological Corporation.

Woodcock, R. W., & Johnson, M. B. (1989). Woodcock-Johnson Psycho-Educational Battery-Revised. Chicago: Riverside Publishing Company.

Woodcock, R. W., & Mather, N. (1989/1990). *Woodcock-Johnson Revised Tests of Achievement: Examiner's manual*. In R. W. Woodcock & M. B. Johnson, Woodcock-Johnson Psycho-Educational Battery–Revised. Chicago: Riverside Publishing Company.

Ysseldyke, J., & Christenson, S. (1993). The Instructional Environment System-II. Longmont, CO: Sopris West.

Lisa Pericola Case received her PhD in Special Education from the University of Maryland in 1997. She is currently a Research Associate at the University of Maryland. Her primary interests are in the areas of definition, diagnosis, and instruction for individuals with learning disabilities.

Deborah L. Speece, PhD, is Professor of Special Education at the University of Maryland. She received the doctorate in educational psychology from the University of North Carolina at Chapel Hill. Her research interests include children at risk for school failure, early identification of reading disabilities, and classification.

Dawn Eddy Molloy, PhD, is Faculty Research Associate and Undergraduate Program Coordinator in the Special Education Department at the University of Maryland. She received her PhD in Special Education at the University of Maryland in 1996. Research interests include children experiencing school failure, teacher development, and qualitative methodologies.