Diagnosing Reading Disability: Reading Recovery as a Component of a Response-to-Intervention Assessment Method

Michael W. Dunn[1]
Washington State University Vancouver

There is growing evidence that the current method of identifying students with a learning disability (LD) is ineffective. The wait-to-fail model of assessing students after second/third grade as well as conceptual problems in using intelligence tests for identification result in students not receiving the assistance they need during the early-elementary school years. The educational community is discussing response to intervention (RTI) as an alternative assessment method. This study explored the assessment components of the Reading Recovery (RR) program as part of a future RTI model. By means of a discriminant-function analysis, a retrospective study of third- through fifth-grade students who participated in RR during first grade investigated assessment elements (beginning text level, ending text level, and number of weeks' participation in RR) of this program. The results showed that RR assessment elements were significant predictors of first-grade students who were later identified as having a reading disability. Ending text level was consistently the largest predictor of students later classified as having a reading disability or not.

Key Words: Reading Disability, Response to Intervention, Predicting RD Status

In considering whether Reading Recovery would be a valuable component of response to intervention (RTI), it is important to understand the nature of RTI and the educational context in which the movement emerged. By means of providing such a context, the following section gives a brief overview of the reasons behind the push for RTI as a way of meeting the needs of students with learning and reading disabilities.

The Reasons for RTI

Advocates for students showing the characteristics of having learning disabilities (LD) petitioned Congress for years before learning disabilities were finally included in the Education for All Handicapped Children Act (PL 94-142; 1975) as a disability category, thereby guaranteeing services to these students. In 1977, the U.S. Department of Education issued a regulation to guide practitioners in defining LD as those students who have a "severe discrepancy" between test scores on performance (IQ) and academic achievement (U.S. Office of Education, 1977). State education departments implemented the discrepancy concept but defined it in different ways, including the difference between a standard IQ score and an academic achieve-

1. Please address correspondence to: Michael W. Dunn, Washington State University Vancouver, Library Building, Room 210-J, 14204 NE Salmon Creek Avenue, Vancouver, WA 98686; E-mail: miwdunn@msn.com

ment score, the regression of IQ on academic achievement, the amount of discrepancy (e.g., 1.0 *SD* vs. 2.0 *SD*s), and the choice of IQ and academic achievement tests. These various methods resulted in large inconsistencies in LD identification rates between states (Scruggs & Mastropieri, 2002).

To address the growing discontent with this assessment and classification method for students as having an LD, the President's Commission on Special Education was convened in 2002 (Fuchs, Mock, Morgan, & Young, 2003). The focus of the discussion centered on two issues. First, there are conceptual problems with the use of intelligence tests in the assessment process for special education services eligibility. Research (i.e., Tal & Siegel, 1996) has found no difference between the reading, spelling, phonological skills, and reading comprehension of individuals with reading disabilities with high IQ scores versus those with low IQ scores. IQ tests do not help predict students who would benefit from remediation (Kershner, 1990). Other research (Stanovich, 1988) indicates that difficulties with reading may impede the development of language, knowledge, and vocabulary skills—the "Matthew effect." This further complicates the relationship between reading and IQ and, therefore, the justification for using IQ in the identification of LD. Second, the practice of waiting until third grade to assess if a student has succeeded in grasping the academic content (the "wait-to-fail" model) is considered to be contributing to the increased severity of academic difficulties for students in older grades (Lyon, Fletcher, Shaywitz, Shaywitz, Torgesen, Wood et al., 2001).

The alternative being proposed and offered to states and districts in the recent reauthorization of Individuals with Disabilities Education Improvement Act (IDEIA) (2004) is RTI (Compton, 2006; Fuchs & Fuchs, 2007; Gresham, 2002; President's Commission on Excellence in Special Education, 2002; Vaughn & Fuchs, 2003). Specifically, a student considered to be struggling with literacy skills may be considered for identification as having an LD by being dually discrepant: low achievement and making little or no progress in a three-tiered intervention program.

In the first tier, students participate in presumably research-based reading instruction activities in the general education classroom. Each student's rate of reading growth is evaluated. A student who is dually discrepant is designated as at risk for poor reading outcomes, and possibly having a reading disability (RD). This student moves to the second tier in the RTI process. Progress monitoring is conducted again—this time in a small-group or individual instructional format referred to as Tier Two. A student at this level of RTI would receive intensive instruction in hopes of demonstrating improved reading development. If progress does occur, the student returns to the general education classroom and is no longer viewed as dually discrepant. If the student does not make good progress in the second tier, an intrinsic deficit is probable, and a need for the third and final RTI tier is evident; a condensed special education evaluation would determine a possible disability classification and placement (Fuchs et al., 2003).

Reading Recovery as an RTI Method

This study investigated RR (Clay, 2002b) as a standard protocol RTI format for LD in reading identification; eighty percent of students identified as having LD have the disability in reading (Roush, 1995). The standard protocol approach of RTI

uses a set of procedures for students with similar characteristics of academic difficulty (Fuchs et al., 2003). For example, one teacher and one student work together on a set of activities for a portion of the school day over a set time period (i.e., 20 weeks) with the aim of improving the student's academic performance. The RR program (Clay, 2002a) closely parallels this format of RTI. With its consistent implementation and instructional methods, the RR program offers the educational community a practical, already in-use component for RTI.

RR, a first-grade remedial literacy program, is implemented annually in over 10,000 U.S. schools as well as in Canada, the United Kingdom, Australia, and New Zealand (Reading Recovery, 2006). RR's daily programming includes students reading texts that require a working knowledge of phonemic awareness, oral fluency, and comprehension. Students who achieve 90% mastery of a text by orally decoding the words advance to the next book level in the series (A, B, 1-30). When students do not make progress through the daily 30-minute literacy lessons tailored for each student per day over 20 weeks, impaired reading skills and a need for further special education services are indicated.

The RR program is designed to address the needs of students experiencing difficulties with literacy (some of whom may later be classified as having a reading disability, RD). The characteristics of RD include difficulties with phonemic awareness (the ability to segment and manipulate the sounds of spoken words), speech perception (the ability to hear subtle differences between words—such as *mat* and *sat*), vocabulary skills, phonetic working memory (successfully repeating a sentence as heard), as well as syntax (i.e., use of a capital letter to start a sentence, combining sentences to make a paragraph) and semantics (i.e., social context of language, idioms) (Mann, 2003). Students experiencing difficulty with one or a combination of these characteristics may have an RD. Relative to these skills, the RR program includes the following activities: oral reading and writing of various genres of text, working with words using plastic letters, solving unknown words by breaking words into parts as well as relating the context of the text to the sentence and unknown word, developing reading fluency, and completing a daily Running Record oral reading assessment; this assessment involves an analysis of students' oral reading miscues and the required cut-off of 90% accuracy which determines if a student is ready to progress to the next book level in the RR program (Clay, 2002b).

RR meets the criteria of RTI with the pass/fail component of students progressing to book level 15 during the 20 weeks of 30-minute, daily reading strategy intervention sessions; the number of sessions can be extended for students who need additional remediation. Due to conceptual problems with the current method of diagnosis for LD used in the wait-to-fail model, as mentioned above, assessment elements of the RR program could serve to predict students who need special education services and to provide them as soon as first grade. Cavanaugh, Kim, Wanzek, and Vaughn (2004) completed a synthesis of 27 intervention studies, which indicated that early intervention for reading difficulties can be effective for students as young as kindergarten.

RR's effectiveness has been challenged in three areas. First, RR is not economical: one teacher per student for 30 minutes per day over 20 weeks (or more). However, since the aim of the RTI model is to provide intensive direct instruction to

address a student's unique needs, an individual academic intervention has merit because the teacher providing the intervention can tailor the level and progression of activities relative to the student's individual readiness level. Also, while classroom-based, problem-solving RTI models do exist, they are not the only acceptable research-based RTI models; standard protocol approaches have also been found to be effective (i.e., Vellutino, Scanlon, Sipay, Small, Pratt, Chen et al., 1996). Furthermore, RR has a high degree of fidelity of treatment, given its extensive training, practice, and ongoing feedback for teachers. This helps ensure that the student receives the intensive intervention (Clay, 2002b). A second criticism is that RR's learning levels may not be sustained in subsequent grades. Finally, 10 to 30% of children receiving the program in first grade (ages 6 to 7) may not successfully complete it (Grossen, Coulter, & Ruggles, 2004; Hiebert & Taylor, 1994). However, the latter two criticisms have been contradicted by research affirming that RR is an effective intervention for students with low literacy skills (Agostino & Murphy, 2004; Brown, Denton, Kelly, Outhred, & McNaught, 1999; Pinnell, 1989; Schwartz, 2005). These research findings were published by highly regarded journals that have no vested interest in supporting RR per se. For example, Agostino and Murphy conducted a meta-analysis of 36 RR studies and found that the program offered positive outcomes for discontinued and non-discontinued students on assessments tailored for the program as well as standardized achievement measures; the meta-analysis indicated a lasting program effect by the end of second grade and beyond.

RR's beginning/ending text levels and set number of weeks for completion could offer educators a means to determine which students would later be identified as RD. Beginning reading skills are an indicator of future reading ability. Bishop (2003) and Catts, Fey, Zhang, and Tomblin (2001) indicated that assessing a student's reading skills as early as kindergarten is a good predictor of students who need assistance with reading skills. Beginning text level relates to RTI's dual discrepancy component in two ways: (a) it helps determine the degree of the student's low performance with reading skills, and (b) it provides a baseline for measuring reading growth over time. In the RR program, book level 15 represents end-of-first-grade reading ability. If a student does not reach this ending text level of ability, he could be defined as having an RD by the end of first grade. Number of weeks' participation in the program relates to a student's ability to make adequate progress with literacy skills during the intervention period. The longer a student "continues" in the program (especially after 20 weeks), the higher the likelihood of an underlying language processing problem (i.e., RD) (Rhodes-Kline, 1996).

Socioeconomic Status and RD

Social class has been identified as a determinant of a student's behavior and performance in school (Grundmann, 1997; O'Connor & Spreen, 1988). Students from lower-income families often experience fewer literacy activities within the home and little opportunity for out-of-school educational experiences. This results in students having less background knowledge and skills that schools demand as a precursor for academic learning. In this study, socioeconomic status was based on students' eligibility for the free/reduced-cost lunch program in first grade.

Contribution of This Study to RTI Research

This study expands previous research in four ways. First, it adds to the body of RTI research in that no other retrospective studies are known to have been completed. Other researchers (i.e., Vaughn, Mathes, Linan-Thompson, & Francis, 2005) have investigated experimental RTI methods or programs using the principles of RTI that only in certain cases aim to identify students. Fuchs et al. (2003) commented on the paucity of studies that analyze intervention programs already in practice. Second, this is the first retrospective study of students who received a reading skills intervention in first grade, some of whom were later identified as having an RD. Third, although Lose et al. (2007) commented on how RR fits the RTI paradigm, this is the first known analysis of RR in an RTI context. Fourth, the results of this study provide a way of determining how the characteristics of RR's book levels and the number of weeks students participate in the program are related to students later being identified as having an LD in reading. Defining the cut-off score is a key component to implementing an RTI model in a school/district. RR provides curriculum-based measures that could help define which students succeeded with the intervention.

Research Questions

This study was designed to address the following questions: Which, if any, of the elements (beginning text level, ending text level, number of weeks of participation in the RR program) are good predictors of students who are later identified as having an RD by third through fifth grade? Do alternative definitions of reading disability (IQ/achievement discrepancy, reading composite scores<30, 23, and 15) indicate RR assessment elements that would be useful in determining RD status under an RTI format? As a covariate to RR assessment components, how does socioeconomic status compare in terms of RD/non-RD status?

Method

Sample

Third- to fifth-grade students (N=155) who participated in RR during first grade formed the basis of the data for the analysis (see Table 1). The sample consisted mostly of white (61%) and black students (30%). Thirty-seven percent of the students represented in the sample were female. Twenty-three percent (35 students) were officially identified as having an LD as defined by participating school districts.

Procedure

Special education personnel (special education directors, Title I directors, and RR teacher leaders) of school districts (A, B, and C) in a midwestern state agreed to participate and provide the required data for the sample (N=155). Of the 35 students with an identified LD, 4 came from School District A, 20 from School District B, and 11 from School District C. The participating school districts identified students as having an LD based on the "wait-to-fail" model of an IQ/achievement discrepancy of 18 points. For students in the sample who were not identified by their districts as having an LD and who had not completed the Wechsler Intelligence Scales for Children (WISC) (Wechsler, 1991), results from the InView (CTB/McGraw Hill, 2002) cognitive abilities test were used to factor out students with possible characteristics of other disabilities (i.e., mental retardation).

Table 1
Descriptive Statistics (N=155)

Gender	
Gender Ratio[1]	M=97/F=58
Age	
Age	
M (*SD*)	9.97 (.764)
Race/Ethnicity	
Asian	0.6%
Black	30.3%
Hispanic	0.6%
Multiracial	2.6%
White	61.3%
Other	4.5%
Grade During 2004-2005 School Year	
Grade 3	13%
Grade 4	43%
Grade 5	45%
Receiving Special Education Services During 2004-2005 School Year	
No	77%
Yes	29%
Retained in a Grade 2	
Yes (1 year)	23%
No	65%
Number of Weeks in Reading Recovery During First Grade	
Mean	16.54
Median	17.00
Mode	20.00
Range	4 to 26
Free/Reduced-Cost Lunch Status in First Grade	
Free/Reduced	85 (55%)
No Subsidy	70 (45%)

[1]M=63%/F=37%.
[2]One school district did not provide data for 25 students' retained-in-grade status.

As a retrospective study, reading composite achievement data on students defined as RD or non-RD in third to fifth grade were analyzed relative to their first-grade RR scores and free/reduced-cost lunch status. School Districts A and B provided TerraNova (TN) reading achievement data (CTB/McGraw Hill, 2001); School District C provided reading achievement results from the Woodcock-Johnson (WJ) III-Tests of Achievement (Woodcock & Johnson, 2001).

Rationale for Analyses

Analysis 1 (RD=35; non-RD=120) used the school districts' definition of LD for defining group membership. Analysis 2 (RD=50; non-RD=105) defined RD/non-RD status by using TN (2001) reading composite scores lower than the 30th percentile. In a similar study, Vellutino et al. (1996) provided an intensive intervention to first- and second-grade students. Students who remained below the 30th percentile on the Woodcock Reading Mastery Test-Revised (WRMT-R) following the intervention were defined as "difficult to remediate." Analysis 3 (RD=35; non-RD=120) used TN scores of less than 23 (23 to 40 represent the slightly below-average range) Analysis 4 (RD=21; non-RD=134) even further refined the definition of RD by using a cut-off reading composite score of 15 (scores of 11 to 22 represent the well-below-average range of the TN test).

Variables Used in the Study

RD/non-RD status. The grouping variable (GV) of the four analyses of the sample (*N=155*) was students' identification as RD or non-RD. The standard procedure for discriminant-function analysis requires that the comparison group size (RD group) be five times the number of variables in the equation (4 variables [beginning text level, ending text level, number of weeks, and free/reduced-cost lunch status] x 5 = 20) (Tabachnick & Fidell, 2001). For this purpose, the RD group consisted of at least 20 students for each analysis.

Assumptions of discriminant-function analysis. All predicting variables (PVs) were within normal limits except for beginning text level. However, it would be expected that beginning text level would not have a normal distribution because students in the RR program would be chosen for not being good readers. Therefore, their beginning text levels would be expected to be at a low book number in the program's series of easy-to-difficult books (book 15 representing the desired ending text level—end of first grade level of ability). For the equality of variance/covariance assumption, the Box's M test for Analysis 1 rendered a significant result; this is attributed to the two groups in this sample (RD and non-RD) being unequal, indicating a violation of the equality of variance/covariance assumption. Similar to the non-normal distribution of beginning text level mentioned earlier, RD/non-RD groups' Box's M tests were significantly different since the two groups represented discrepant levels of IQ and academic achievement. The results for Analysis 2, 3, and 4 indicated a nonsignificant Box's M test, supporting the robustness of these analyses. The multi-collinearity assumption (low correlations between PVs) was met (Tabachnick & Fidell, 2001).

Beginning/ending text levels. These PVs were continuous variables; book levels were coded A, B (pre-kindergarten) and 1-30. Book levels A and B contain one word per page, indicating an object that changes color as the pages progress. Book

level 5 is considered representative of the beginning of first grade. Book level 15 (end of first grade) represents a more varied and challenging form of text. For example, the place of text varies throughout the story (top, middle, bottom of page), some pages have one sentence, other pages have three, and there are changes in verb tenses as well as types of sentences (interrogative, imperative). Book level 30 would be similar to an early third-grade chapter book series (Clay, 2002b).

Number of weeks. Number of weeks in the RR program can vary from one district to another; for the school districts that participated in this study, 20 weeks was considered standard. Students who attained book level 15 before or at week 20 were considered as "discontinued" (successful). Students who did not reach book level 15 were considered "continued" (unsuccessful). These students may have received additional sessions to improve reading skills. This PV was a continuous variable (1-20 or more).

Free/reduced-cost lunch status. Free/reduced-cost lunch status was used as a means of categorizing the socioeconomic status of students in the sample. This categorical PV was coded as either 0 (not eligible for any subsidy) or 1 (reduced-cost or free lunch). The free/reduced-cost lunch data were based on the student's family income during his or her participation in the RR program in first grade.

Children may participate in the free lunch program in two ways. First, the household is a participant in Food Stamps, Temporary Assistance for Needy Families, or the Food Distribution Program on Indian Reservations. Second, the student's household income falls below 130% of the federal poverty level. For reduced-priced meals, household income is between 130 and 185% of the federal poverty level ("National School Lunch Program," n.d.).

Methods of Analysis

The aim of this study was to examine whether a relationship exists between RR assessment scores and a student's subsequent identification as RD. A discriminant-function analysis was completed using SPSS (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975). Beginning text level, ending text level, number of weeks in the RR program, and free/reduced-cost lunch status (as a covariate) represented the PVs of the function. Group membership (RD or non-RD) served as the GV.

Four student cases from which data for the GV and PVs were missing were deleted from the sample. Since assessment scores between RD and non-RD groups can vary widely, imputing values for missing data could change groups' mean scores and, therefore, possibly render results unreflective of the two groups. In the resulting analyses, coefficients above +/-.500 were interpreted. Standardized discriminant-function coefficients compare the relative importance of the PVs. Structure coefficients are the correlations between a given PV and the discriminant scores associated with the discriminant function; structure coefficients indicate how closely a variable is related to each function. A table/list of structure coefficients is referred to as a structure matrix. To assess for power in the discriminant-function analyses, eigenvalues, which indicate relative discriminating power, were used (Cohen, 1988; Discriminant Function Analysis, n.d.; Tabachnick & Fidell, 2001). Independent samples *t* tests were conducted to evaluate if there were statistically significant differences between the two groups in each analysis of this study (RD and non-RD) based on beginning text level, ending text level, number of weeks, reading composite, and IQ variables.

Results

Results of Discriminant Functions

Tables 2-4 show the results of the analyses. The correlational matrix indicated that no correlations resulted in a value beyond +/-0.900. This suggests that multicollinearity was not an issue in the discriminant function. The correlation between IQ and reading composite was moderate (r=.494, p<.01). Number of weeks was negatively correlated with beginning text level (r=-.428, p<.05). The Wilks's lambda for each of the four analyses were significant (Analysis 1: Λ = .828, X^2[4, N =155] = 28.58, p <.001; Analysis 2: Λ = .854, X^2[4, N =155] = 23.82, p <.001; Analysis 3: Λ = .907, X^2[4, N =155] = 14.69, p <.005; Analysis 4: Λ = .930, X^2[4, N =155] = 10.88, p <.028) indicating that the predictors differentiated among the two student groups (RD/non-RD). Eigenvalues indicated the analyses' relative discriminating power: Analysis 1=.21, Analysis 2=.17, Analysis 3, .10, and Analysis 4=.08. The standardized discriminant-function coefficients and structure coefficients are presented in Table 3.

Table 2
Intercorrelational Matrix (N = 155)

	1	2	3	4	5	6	7	8
1. Age	—	-.215**	-.019	.000	-.301*	-.073	.020	.191*
2. IQ[1]		—	494**	-.002	.075	.247*	-.135	.007
3. Reading Composite[2]			—	-.035	.007	.241*	.050	.256**
4. RR Number of Weeks				—	-.428*	-.214*	076	149
5. Beginning Text Level[3]					—	150	-.043	-.061
6. Ending Text Level[4]						—	-.250*	-.342**
7. Free/Reduced-Cost Lunch[5]							—	.274**
8. RD Status[6]								—

*Correlation is significant at the 0.05 level (2-tailed).
**Correlation is significant at the 0.01 level (2-tailed).

[1] InView/WISC III IQ Test scores.
[2] Terra Nova Academic Achievement Test/Woodcock-Johnson III Tests of Achievement scores.
[3] RR instruction and assessment book series.
[4] RR instruction and assessment book series.
[5] Students' free/reduced-cost lunch status during first grade.
[6] Student's RD or non-RD status as of the 2004-2005 school year.

Table 3

Discriminant-Function Analysis to Classify Correctly RD vs. Non-RD Group Membership

	Wilks's Lambda	Standardized Discriminant-Function Coefficients[1]		Structure Coefficients[1]		Group Centroids
Analysis 1		ETL[2]	.783[6]	ETL	.890	
School District Definition of RD	.828 (p <.001)	F/RL[3]	-.404	F/RL	-.547	
RD (N=35) and		Weeks[4]	-.239	Weeks	-.394	RD: -.840
Non-RD group (N=120)		BTL[5]	-.079	BTL	.146	Non-RD: .245
Analysis 2		ETL	1.017	ETL	921	
Reading Composite < 30	.854 (p <.001)	F/RL	380	Weeks	-.262	
RD (N=50) and		BTL	-.152	BTL	049	RD: -.595
Non-RD group (N=105)		Weeks	-.130	F/RL	095	Non-RD: .283
Analysis 3						
Reading Composite <23		ETL	1.054	ETL	967	
RD (N=35) and	.907 (p <.005)	F/RL	163	Weeks	-.102	
Non-RD Group (N=120)		BTL	-.144	F/RL	-.078	RD: -.588
		Weeks	092	BTL	-.024	Non-RD: .172
Analysis 4						
Reading Composite <15		ETL	.957	ETL	855	
RD (N =21) and	.930 (p <.028)	Weeks	.531	F/RL	-.346	
Non-RD Group (N=134)		F/RL	.186	Weeks	255	RD: -.686
		BTL	.033	BTL	-.053	Non-RD: .107

[1]Values for each analysis are listed in descending order.
[2]Ending text level in the RR program.
[3]Free/reduced-cost lunch status during first grade.
[4]Number of weeks in the RR program.
[5]Beginning text level in the RR program.
[6]The higher the ending text level, the less likely to be identified as having an RD.

Table 4

Descriptive Statistics of Reading Composite, Number of Weeks in RR, Beginning Text Level, Ending Text Level, and Free/Reduced-Cost Lunch Status in First Grade

	Analysis 1 School Districts' Definition		Analysis 2 Reading Composite<30		Analysis 3 Reading Composite<23		Analysis 4 Reading Composite<15	
	RD (*N*=35)	Non-RD (*N*=120)	RD (*N*=50)	Non-RD (*N*=105)	RD (*N*=35)	Non-RD (*N*=120)	RD (*N*=21)	Non-RD (*N*=134)
Reading Composite Mean[1] (*SD*)	54.40 (28.891)	42.71 (22.767)	16.36 (8.086)	59.15 (16.566)	12.23 (5.806)	55.01 (19.016)	8.62 (4.341)	51.10 (21.349)
	Cohen's *d*=.51[2]		Cohen's *d*=-2.58		Cohen's *d*=-2.25		Cohen's *d*=-1.99	
Number of Weeks in RR Mean[3] (*SD*)	17.97 (3.97)	16.12 (4.44)	17.22 (4.097)	16.21 (4.510)	16.80 (4.276)	16.46 (4.442)	15.86 (4.127)	16.64 (4.439)
	Cohen's *d*=0.42		Cohen's *d*=0.24		Cohen's *d*=0.07		Cohen's *d*=-0.18	
Beginning Text Level Mean (*SD*)	0.89 (1.83)	1.28 (2.67)	1.12 (2.076)	1.23 (2.694)	1.23 (2.211)	1.18 (2.593)	1.29 (2.194)	1.18 (2.557)
	Cohen's *d*=-0.15		Cohen's *d*=-0.04		Cohen's *d*=0.02		Cohen's *d*=0.04	
Ending Text Level Mean[1] (*SD*)	8.40 (4.97)	14.57 (6.74)	9.64 (6.009)	14.86 (6.640)	9.43 (5.700)	14.27 (6.822)	9.24 (6.625)	13.79 (6.729)
	Cohen's *d*=-0.90		Cohen's *d*=-0.80		Cohen's *d*=-0.71		Cohen's *d*=-0.68	
Free/Reduced Lunch Frequency (%)	27 (77%)[4]	58 (48%)	26 (52%)	59 (56%)	20 (57%)	65 (54%)	14 (67%)	71 (53%)

[1]Independent samples *t* tests indicated significant differences for RD and non-RD groups in all four analyses.

[2]Estimate of the effect size using Cohen's *d* (mean of RD – mean of non-RD/SD of the control group [non-RD group]).

[3]Independent samples *t* tests indicated significant differences for RD and non-RD groups in Analyses 1 and 2.

[4]For example, in Analysis 1, 27 of 35 students classified as RD were on free/reduced-cost lunch programs in first grade (77% of the RD group).

Ending text level was consistently the largest PV in the four functions and was most representative of the functions as defined by the structure matrix; free/reduced-cost lunch status was well represented in the function of Analysis 1. Number of weeks became a significant PV only in Analysis 4 (reading composite <15). Each successive analysis (1-4) resulted in explaining less and less of the variance in the function of the four variables.

Table 4 indicates that in Analysis 1 (school districts' definition), reading composite was higher for the RD group than the non-RD group. In Analyses 2-4, students categorized as RD demonstrated lower levels of functioning as indicated by their reading composite and ending text-level scores. Beginning text level followed this pattern except in Analyses 3 and 4 where the RD group had a marginally higher value. With the exception of Analysis 2, the RD group had comparatively larger number of students who participated in free/reduced-cost lunch programs. The number of weeks was relatively constant across RD/non-RD groups in all four analyses.

Discussion

By means of a discriminant-function analysis, this research study examined whether beginning text level, ending text level, number of weeks in the RR program, and free/reduced-cost lunch status (as a covariate) would be good predictors of students identified as having an RD by third through fifth grade. The results indicated that higher ending text level was the largest PV of each of the four analyses; there was also a small but significant negative correlation with ending text level; that is, the higher the ending text level, the less likely a student is to be identified as having an RD. Number of weeks was a good PV in Analysis 4 (reading composite <15). The negative correlation with beginning text level would be expected; the lower the initial text level, the more likely a student would need to complete a relatively higher number of weeks in the RR program. The slightly higher beginning text levels of Analyses 3 and 4 mean value would still round to book level 1. The mean number of weeks did not vary much between groups across the four analyses. This may be attributed to such factors as some students' parents requesting a withdrawal from the program, students being absent from school, or student relocating to another school district. With 55% of the students in this study on free/reduced-cost lunch, a large portion of the sample was from low-income families who may be transient or truant (Stronge, 1992). Discriminating power as demonstrated by the eigen-values was relatively lower in Analysis 4 at least in part due to a smaller RD-group sample size; as sample size decreases, the less it reflects the population it aims to represent (Martella, Nelson, & Marchand-Martella, 1999).

Research (Grundman, 1997; O'Connor & Spreen, 1988) has discussed a link between socioeconomic status and LD. In terms of descriptive statistics, the sample in this study represents an RD-low socioeconomic relationship, given that 55% of the sample was on free/reduced-cost lunch programs. Although free/reduced-cost lunch status was well represented in the discriminant function of Analysis 1 (traditional IQ/achievement discrepancy method), it was not a significant predictor in any of the four analyses. With an assessment of intelligence not being part of the assessment models of Analyses 2-4, socioeconomic as well as racial and

cultural biases inherent in IQ testing would be evaded (Coutinho, Oswald, & Best, 2002).

Many researchers (e.g., Peterson & Shinn, 2002; Warner, Dede, Garvan, & Conway, 2002; Watkins, Kush, & Schaefer, 2002) argue against continued use of IQ in the identification of RD. The results of the present study supported this argument in that IQ and RD were not highly correlated. Although the IQ/achievement identification model in Analysis 1 accounted for relatively more variance than the more refined definitions of RD based on reading achievement scores, the difference between them was negligible. This may be attributed to the successively smaller RD group sizes in Analyses 2-4. That is, as the size of the comparison group declines, the power of the analysis is impacted. Furthermore, the moderate correlation between IQ and reading composite helps support the use of reading composite as a replacement for IQ in classifying students as RD/non-RD in the analyses. Also, the school districts' definition resulted in a higher mean reading composite score for the RD group than the non-RD group; Analysis 2-4 resulted in the more rational result.

Findings Relative to Other RTI Research

The RR program meets the criteria used in RTI research (Fuchs & Fuchs, 1998) with its dually discrepant method of assessment (progress through leveled texts during the 20 weeks of remediation) as well as the pass/fail component of reaching book 15 (representative of first-grade level of reading ability) after 20 weeks of participation in the intervention. This model reflects the objectives of RTI's Level 2: (a) to prevent reading difficulty by delivering an intensive, and presumably effective, intervention that improves reading development; and (b) to assess the level of responsiveness to an instructional intensity from which most students' performance should improve.

The results of this study indicate that ending text level is a significant indicator of students who would later be identified as having an LD, 80% of which have an RD (Roush, 1995). School districts that use RR should incorporate ending text level into their identification practices when considering RD status. It would not explain the entire concept of having an RD, but it would be an indicator of the need for further assessment. In addition, the students with low reading skills could be referred by the end of first grade as opposed to waiting until third or fourth grade.

Although the increasingly refined definitions of RD in this study rendered significant functions and consistently rated ending text level as the largest discriminating variable, the declining amount of variance explained would suggest the presence of other factors in determining LD status. Kavale, Holdnack, and Mostert (2006) commented that an RTI model cannot stand alone as the primary means of identifying for LD. The results of the present study would support this, given that ending text level only explains 7 to 17% of the variance in the construct of LD identification. That is, RR's ending text level alone would not be sufficient as an RTI model. Other aspects related to assessment for LD need to be considered.

More explicit assessments of the characteristics of RD (phonemic awareness, speech perception, vocabulary skills, phonetic working memory, syntax, and semantics) during an intervention program such as RR could better predict students to be considered for RD classification. For example, other research (Lovett, Steinbach, & Fritjers, 2000) has found explicit assessments of rapid automatic nam-

ing (RAN) and phonemic awareness to be good predictors of students having difficulty with reading skills. Periodic evaluations of these aspects of reading skills in addition to daily RR programming could help provide a more comprehensive picture of students' progress based on the dual discrepant model.

Limitations

The present study did not measure the type and quality of general education classroom instruction that the students received before, during, and after their participation in the RR program. Students may or may not progress with RR activities due to the type or depth of literacy instruction that occurs within the classroom. Although research (Roush, 1995) has found that about 80% of students with an LD have low skills in the area of reading, access to individual students' files was not included for the data set used here. Nevertheless, students' participation in the RR program due to difficulty with literacy tasks would suggest that at least 80% had difficulties related to reading.

Generalization to the larger national student population is hindered by regional and demographic factors. The sample for this study was composed of students from three school districts in a midwestern state. The proportion of racial groups in the sample is not representative of students across the nation. For example, the study included 30% black students, whereas black students represent 14.8% of the national student population. Further, 1% of the sample was Hispanic compared to 14.2% of the American school population (Lawson et al., 2002). RR has been funded through government programs and legislation (i.e., No Child Left Behind Act of 2001) to address the needs of students of low socioeconomic status. Students with low SES background represented 55% of the sample in this study, whereas about 18% of U.S. children were living in poverty as of 2004 (Koball & Douglas-Hall, 2006).

Future Research

As a conceptual model of identification for students with LD generally, RTI needs to define not only the cut-off score to be used but also other elements of the assessment-for-identification process. For RD, this study highlighted the aspect of ending text level as a significant predictor in all four analyses; however, only 7 to 17% of the variance in RD/non-RD group membership was explained by ending text level, beginning text level, number of weeks in RR, and free/reduced-cost lunch status. Thus, other components must be considered in the identification of RD.

Future research could investigate whether combining RR with the model proposed by Flanagan, Ortiz, Alfonso, and Mascolo (2002) would provide a more comprehensive RTI assessment process. Flanagan et al.'s model is a comprehensive assessment that includes quantitative knowledge (math calculation and reasoning), reading and writing (oral reading, comprehension, written expression), and crystallized intelligence (general information, oral expression, lexical knowledge, and listening comprehension). Combining the two models (splitting level 2 of the RTI format into two steps) would help define the factors that could account for the remaining variance not explained in the discriminant functions of this study and help schools' multidisciplinary teams decide possible classification for long-term special education services: level 3 of RTI.

References

Agostino, J. V., & Murphy, J. A. (2004). A meta-analysis of Reading Recovery in United States schools. *Educational Evaluation and Policy Analysis, 26*(1), 26-38.

Bishop, A. (2003). Prediction of first-grade reading achievement: A comparison of fall and winter kindergarten screenings. *Learning Disability Quarterly, 26*(3), 189-200.

Brown, W., Denton, E., Kelly, L., Outhred, L., & McNaught, M. (1999). RR's effectiveness: A five-year success story in San Luis Coastal Unified School District. *ERS Spectrum, Winter*, 3-10.

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (2001). Estimating the risk of future reading difficulties in kindergarten children: A research-based model and its clinical implementation. *Language, Speech, and Hearing Services in Schools, 32*(1), 38-50.

Cavanaugh, C. L., Ae-Hwa Kim, C. L., Wanzek, J., & Vaughn, S. (2004). Kindergarten reading interventions for at-risk students: Twenty years of research. *Learning Disabilities—A Contemporary Journal, 2*(1), 9-21.

Clay, M. (2002a). *An observation survey of early literacy achievement.* Portsmouth, NH: Heinemann.

Clay, M. (2002b). *Reading Recovery: A guidebook for teachers in training.* Portsmouth, NH: Heinemann.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Compton, D. L. (2006). How should "unresponsiveness" to secondary intervention be operationalized? It's all about the nudge. *Journal of Learning Disabilities, 39*(2), 170-173.

Coutinho, M., Oswald, D., & Best, A. (2002). The influence of sociodemographics on the disproportionate identification of minority students as having learning disabilities. *Remedial and Special Education, 23*(1), 49-59.

CTB/McGraw Hill. (2002). *InView Cognitive Abilities Test.* Monterey, CA: Author.

CTB/McGraw Hill. (2001). *TerraNova, The second edition.* Monterey, CA: Author.

Discriminant Function Analysis. (n.d.). *Overview.* Retrieved May 21, 2007, from http://www2.chass.ncsu.edu/garson/pa765/discrim.htm

Education for All Handicapped Children Act. (1975). Pub. L. 94-142, 20 U.S.C. § 1401.

Flanagan, D., Ortiz, S., Alfonso, V., & Mascolo, J. (2002). *The achievement test desk reference (ATDR): Comprehensive assessment and learning disabilities.* Boston: Allyn & Bacon.

Fuchs, D., & Fuchs, L. (2007). Responsiveness to intervention [Special issue]. *Teaching Exceptional Children, 39*(5).

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

Fuchs, D., Mock, D., Morgan, P., & Young, C. (2003). Responsiveness-to-instruction: Definitions, evidence, and implications for learning disabilities construct. *Learning Disabilities Research & Practice, 18*(3), 157-171.

Gresham, F. (2002) Responsiveness to intervention: An alternative approach to the identification of learning disabilities. In R. Bradley, L. Danielson, & D. Hallahan (Eds.), *Identification of learning disabilities: Response to treatment* (pp. 467-519). Mahwah, NJ: Erlbaum.

Grossen, B., Coulter, G., & Ruggles, B. (2004). *Reading Recovery: An evaluation of benefits and costs.* Retrieved November 14, 2006, from http://darkwing.uoregon.edu/~bgrossen/rr.htm

Grundmann, M. (1997, April). *The linkage between social structure, competence and academic achievement: Results of a longitudinal study from ages 7 through 21.* Paper presented at the Biennial Meeting of the Society for Research in Child Development, Washington, DC.

Hiebert, E., & Taylor, B. (1994). *Getting reading right from the start: Effective early literacy intervention.* Boston: Allyn & Bacon.

Individuals with Disabilities Education Improvement Act of 2004, Pub. L. No. 108-446, 118 Stat. 2647 (2004).

Kavale, K., Holdnack, J., & Mostert, M. (2006). Responsiveness to intervention and the identification of specific learning disability: A critique and alternative proposal. *Learning Disability Quarterly, 28*, 2-16.

Kershner, J. R. (1990). Self-concept and IQ as predictors of remedial success in children with learning disabilities. *Journal of Learning Disabilities, 23*, 368-374.

Koball, H., & Douglas-Hall, A. (2006). *The new poor: Regional trends in child poverty since 2000.* New York: National Center for Children in Poverty. Retrieved May 22, 2007, from http://www.nccp.org/publications/pub_672.html

Lawson, Q., Humphrey, L., Wood-Garnett, S., Fearn, K., Welch, C., Greene-Bryant, B., & Avoké, S. (2002). *Addressing over-representation of African-American students in special education.* Washington, DC: Council for Exceptional Children.

Lose, M. K., Schmitt, M. C., Gómez-Bellengé, F. X., Jones, N. K., Honchell, B. A., & Askew, B. J. (2007). *Reading Recovery and IDEA legislation: Early intervening service (EIS) and response to intervention (RTI).* Retrieved May 17, 2007, from http://readingrecovery.org/pdfs/SPEDbrief07.pdf.

Lovett, M. W., Steinbach, K. A., & Fritjers, J. C. (2000). Remediating the core deficits of developmental reading disability: A double-deficit perspective. *Journal of Learning Disabilities, 33*, 334-358.

Lyon, G. R., Fletcher, J. M., Shaywitz, S. E., Shaywitz, B. A., Torgesen, J. K., Wood, F. B., Schulte, A., & Olson, R. (2001). Rethinking learning disabilities. In C. E. Finn, C. R. Hokanson, & A. J. Rotherham (Eds.), *Rethinking special education for a new century* (pp. 259-287). Washington, DC: Thomas B. Fordham Foundation.

Mann, V. (2003). Language processes: Keys to reading disability. In H. Swanson, K. Harris, & S. Graham (Ed.), *Handbook of learning disabilities* (Vol. 1, pp. 213–228). New York: The Guilford Press.

Martella, R. C., Nelson, R., & Marchand-Martella, N. E. (1999). *Research methods: Learning to become a critical research consumer.* Boston: Allyn & Bacon.

National School Lunch Program. (n.d.). Retrieved November 14, 2006, from http://www.frac.org/html/federal_food_programs/programs/nslp.html

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. (1975). *Statistical package for the social sciences* (2nd ed.). New York: SPSS, Inc.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).

O'Connor, S., & Spreen, O. (1988). The relationship between parents' socioeconomic status and education level, and adult occupational and educational achievement of children with learning disabilities. *Journal of Learning Disabilities, 21*(3), 148–153.

Peterson, K.M.H., & Shinn, M. R. (2002). Severe discrepancy models: Which best explains school identification practices for learning disabilities. *School Psychology Review, 31*(4), 459-476.

Pinnell, G. (1989). Reading recovery: Helping at risk children to read. *The Elementary School Journal, 90*(2), 159-181.

President's Commission on Excellence on Special Education. (2002). *A new era: Revitalizing special education for children and their families.* Retrieved May 18, 2007, from http://www.ed.gov/inits/commissionsboards/whspecialeducation/reports/index.html

Reading Recovery. (2006). Retrieved May 22, 2007, from http://www.readingrecovery.org/sections/reading/index.asp

Rhodes-Kline, A. (1996). First and second round reading recovery: What difference does it make for discontinuation and program length? *Literacy, Teaching and Learning, 2*(2), 93-102.

Roush, W. (1995). Arguing over why Johnny can't read. *Science, 267*, 1896-1898.

Schwartz, R. M. (2005). Literacy learning of at-risk first-grade students in the Reading Recovery early intervention. *Journal of Educational Psychology*, *97*(2), 257-267.

Scruggs, T. E., & Mastropieri, M. A. (2002). On babies and bathwater: Addressing the problems of identification of learning disabilities. *Learning Disability Quarterly, 25*(3), 155-168.

Stanovich, K. E. (1988). Explaining the differences between the dyslexic and garden variety poor reader: The phonological-core variance-difference model. *Journal of Learning Disabilities, 21*, 590-604.

Stronge, J. H. (1992). The background: History and problems for schooling of the homeless. In J. H. Stronge (Ed.), *Educating homeless children and adolescents: Evaluating policy and practice* (pp. 3-25). Newbury Park, CA: Sage Publications.

Tabachnick, B., & Fidell, L. (2001). *Using multivariate statistics.* Needham Heights, MA: Allyn & Bacon.

Tal, N. F., & Siegel, L. S. (1996). Pseudoword reading errors of poor, dyslexic and normally achieving readers on multisyllable pseudowords. *Applied Psycholinguistics, 17*, 215-232.

U.S. Office of Education. (1977, December 29). Assistance to states for education of handicapped children: Procedures for evaluating specific learning disabilities. *Federal Register, 42*(250), 65082-65085. Washington, DC: U.S. Government Printing Office.

Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice, 18*(3), 137-146.

Vaughn, S., Mathes, P., Linan-Thompson, S., & Francis, D. (2005). Teaching English language learners at risk for reading disabilities to read: Putting research into practice. *Learning Disabilities Research & Practice, 20*(1), 58-67.

Vellutino, F., Scanlon, D., Sipay, E., Small, S., Pratt, S., Chen, R., et al. (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*(4), 601-638.

Warner, T., Dede, D., Garvan, C., & Conway, T. (2002). One size still does not fit all in specific learning disability assessment across ethnic groups. *Journal of Learning Disabilities, 35*(6), 500-508.

Watkins, M. L., Kush, J. C., & Schaefer, B. A. (2002). Diagnostic utility of the learning disability index. *Journal of Learning Disabilities, 35*(2), 98-103.

Wechsler, D. (1991). *Wechsler intelligence scale for children–third edition: Manual.* San Antonio, TX: The Psychological Corporation.

Woodcock, R., & Johnson, M. (2001). *Woodcock-Johnson III tests of achievement* Itasca, IL: Riverside Publishing.

Submitted: October 10, 2006
Revised: February 11, 2007
Accepted: May 18, 2007