Introduction to Response to Intervention: What, why, and how valid is it?

DOUGLAS FUCHS, LYNN S. FUCHS
Peabody College of Vanderbilt University, Nashville, Tennessee, USA

doi:10.1598/RRQ.41.1.4

On December 3, 2004, President Bush signed into law the Individuals with Disabilities Education Improvement Act (IDEA, 2004). The revised law is different from the previous version in at least one important respect. Whereas practitioners were previously encouraged to use IQ–achievement discrepancy to identify children with learning disabilities (LD), they now may use "Response to Intervention," or RTI, a new, alternative method. It is also a means of providing early intervention to all children at risk for school failure. IDEA 2004 permits districts to use as much as 15% of their special education monies to fund early intervention activities. All this has implications for the number and type of children identified, the kinds of educational services provided, and who delivers them. RTI may be especially important to *Reading Research Quarterly* readers because roughly 80% of those with an LD label have been described as reading disabled (Lyon, 1995). With RTI, there may be a larger role for reading specialists, which in turn might affect pre- and inservice professional development activities conducted by universities and school districts. Yet much still needs to be understood to ensure that RTI implementation will promote effective early intervention and represents a valid means of LD identification. In this article, we explain important features of RTI, why it has been promoted as a substitute for IQ–achievement discrepancy, and what remains to be understood before it may be seen as a valid means of LD identification.

What is RTI?

Explaining the R *in RTI*

Selecting at-risk students

Before data are gathered to determine if students are responsive to (or benefiting from) intervention, a subgroup of at-risk students is identified from which nonresponders are likely to emerge. Identification of this subgroup usually occurs in the first month of the school year. Practitioners may choose among several strategies to accomplish this. They can look at all students' performance on last year's high-stakes test and choose a criterion such as scores below the 25th percentile to designate risk. Alternatively, in the current school year, they may test all students in a given grade and designate those scoring below the same percentile (for a norm-referenced measure) or below a performance benchmark (for a criterion-referenced measure) as at risk. From a measurement perspective, perhaps the best strategy is to assess every student in the grade on a screening tool with a benchmark that demonstrates utility for predicting end-of-year performance on high-stakes tests (elementary grades) or on local graduation requirements (secondary level).

Monitoring at-risk students

Once at-risk students are selected, their responsiveness to general education instruction is monitored. Again, there is more than one way of

doing this. At the end of a relatively short period (e.g., eight weeks) of classroom instruction, at-risk students may be administered a brief standardized achievement test in the area of risk. Responsiveness may be defined as "a score above the 16th percentile." An arguably better method would require practitioners to compare the at-risk students' performance to (a) local or national normative estimates for weekly improvement or (b) criterion-referenced figures for weekly improvement. If (a) and (b) are unavailable, then responsiveness may be operationalized as "some improvement" (i.e., increased achievement greater than the standard error of estimate). At-risk children unresponsive to classroom instruction are given more intensive instruction at a second tier, or level, either in or outside the classroom, as discussed subsequently in this article. And their performance during this more intensive, second-tier instruction may be assessed in a manner similar to how performance was assessed during first-tier instruction provided to all students.

Much of RTI assessment, therefore, is progress monitoring. It is a form of *dynamic* assessment because its metric is change in students' level or rate of learning. Such information assists practitioners' efforts both to design early intervention and to identify special-needs children. Regarding early intervention, progress monitoring can be understood in part as formative evaluation: Teachers use the data to determine whether they need to change their curricula, materials, or instructional procedures. Progress monitoring also generates diagnostic information that helps practitioners make classification and program placement decisions (e.g., moving a student from tier 1 to tier 2).

Explaining the I *in RTI*

Focus on reading instruction

Most educators look to RTI as a means of delivering early intervention to address academic problems, not school behavior problems. Specifically, the interventions typically target reading problems and, more specifically, early reading problems (e.g., Al Otaiba & Fuchs, in press; McMaster, Fuchs, Fuchs, & Compton, 2005; O'Connor, 2000; Vaughn, Linan-Thompson, & Hickman, 2003; Vellutino et al., 1996). This is not accidental. Many of the same policymakers behind RTI were also responsible for Reading First, a major component of No Child Left Behind (2002), which requires schools to use scientific knowledge to guide selection of core curricula and to use valid screening measures and progress monitoring to identify students in need of more intensive instruction. In a sense, RTI may be understood as an important aspect of Reading First and current educational policy.

Multitiered instruction

RTI is also multitiered. Different RTI versions have two to four tiers of instruction (see Fuchs, Mock, Morgan, & Young, 2003). The nature of the academic intervention changes at each tier, becoming more intensive as a student moves across the tiers. Increasing intensity is achieved by (a) using more teacher-centered, systematic, and explicit (e.g., scripted) instruction; (b) conducting it more frequently; (c) adding to its duration; (d) creating smaller and more homogenous student groupings; or (e) relying on instructors with greater expertise. Some practitioners (e.g., Grimes, 2002) regard these tiers as substituting for the comprehensive evaluation now afforded all children suspected of having LD. Others (e.g., Division of Learning Disabilities of the Council for Exceptional Children, n.d.; Telzrow, McNamara, & Hollinger, 2000) see the RTI tiers as a component of a more comprehensive and traditional evaluation. The first group views RTI mostly in terms of providing prevention and advocates for more tiers. The second group regards RTI mostly as an identification and classification procedure and argues for fewer tiers.

Problem solving

To date, practitioners conducting RTI use a problem-solving approach to intervention. Researchers, by contrast, favor the use of standard treatment protocols. To explain problem solving, we turn to the work of practitioners in the Heartland (Iowa) Educational Agency. As part of statewide reform, Heartland staff developed a four-level problem-solving model partly to "provide educational assistance in a timely manner" (Grimes, 2002, p. 8). According to Ikeda and Gustafson (2002), at Level 1, a teacher confers with a student's parent(s) to try to resolve academic or behavior problems. At Level 2, a teacher and his or her school's Building Assistance Team meet to identify and analyze problems and to help the teacher select, implement, and monitor the effectiveness of an intervention. An absence of success at this level triggers the involvement of Heartland staff, which defines Level 3. Heartland staff is mostly doctoral- or master's-level school psychologists and special educators who use behavioral problem solving to refine or redesign the interven-

tion and coordinate its implementation. Finally, at Level 4, special education assistance and due process protections are considered.

At each problem-solving level, the process is meant to be the same: Practitioners determine the magnitude of the problem, analyze its causes, design a goal-directed intervention, conduct it as planned, monitor student progress, modify the intervention as needed (i.e., based on student responsiveness), and evaluate its effectiveness and plot future actions (cf. Grimes, 2002). Throughout this problem-solving process, and across the four tiers, "data about a student's responsiveness to intervention becomes the driving force" (Grimes, p. 4). Teachers and Heartland staff are directed to compare a student's performance level and learning rate with what is expected of other students in the same classroom. It is the student's relative classroom performance, rather than test performance, that determines responsiveness and eventually special education eligibility.

The problem-solving approach to intervention has been adopted by a significant number of school districts (A. Canter, personal communication, February 4, 2003), including the Minneapolis Public Schools (see Marston, Muyskens, Lau, & Canter, 2003). Its popularity among practitioners is no doubt due in part to its idiopathic nature: For each child, an effort is made to personalize assessment and intervention. But this individualized approach is a potential weakness as well as a strength. The problem-solving approach presupposes considerable expertise among practitioners in assessment and intervention. They must be skillful in numerous types of assessment and intervention; they must have the clinical judgment and experience to know which assessments and interventions to apply; and they must have the knowledge, discipline, and opportunity to accurately measure the effectiveness of interventions, which are sometimes a unique hybrid of two or more evidence-based practices that in combination have no track record (see Fuchs et al., 2003, for a review of the evidence on the problem-solving approach).

Standard treatment protocol

A standard treatment protocol is an alternative to problem solving. Whereas the problem-solving approach differs from child to child, a standard treatment protocol does not. Implementation usually involves a trial of fixed duration (e.g., 10 to 15 weeks) delivered in small groups or individually (e.g., Al Otaiba & Fuchs, in press; McMaster et al., 2005; Vaughn et al., 2003; Vellutino et al., 1996). If students respond to the treatment trial, they are seen as remediated and disability-free and are returned to the classroom for instruction. If they are unresponsive, they move to a more intensive, Tier 2 standard treatment protocol. If they then demonstrate adequate progress, they are returned to the classroom. But if they show *insufficient* progress at Tier 2, a disability is suspected and further evaluation is warranted.

This approach is illustrated by the work of Vellutino and colleagues (e.g., Vellutino et al., 1996), who asked first-grade teachers to nominate their poorest readers at the beginning of the school year. At the start of the second semester, Vellutino et al. assigned the children to tutoring and contrast groups. The tutored children received a 30-minute, one-to-one intervention five days each week for most of the semester. This intervention amounted to between 70 and 80 tutoring sessions, which focused on phonemic awareness, decoding, sight-word practice, comprehension strategies, and reading connected text. In the fall of second grade, tutored students below the 40th percentile on the Basic Skills Cluster participated in another eight to ten weeks of tutoring.

Two thirds of the tutored students demonstrated "good growth" or "very good growth" after one semester of first-grade tutoring. Indeed, they had basically caught up to their classmates. Vellutino et al. (1996) suggested that these students had not really been reading disabled but "instructionally" disabled. By contrast, the remaining one third of the tutored readers remained in the lowest 30th percentile on the Word Identification and Word Attack subtests of the Woodcock Reading Mastery Test–Revised (WRMT–R), despite receiving tutoring in both first and second grade. The researchers described these children as "difficult to remediate."

Intervention-as-test

As with assessment, intervention—be it problem solving or a standard treatment protocol—serves RTI's two purposes: to provide struggling students with early, effective instruction and to provide a valid means of assessing learner needs. The *I* in RTI has, in a sense, become the test stimulus. Children's level or rate of growth—their degree of responsiveness—is the test performance. Although many RTI proponents are critical of the traditional psychometric approach, they still must prove the validity of their methods; in this case, intervention-as-test. A principal means of demonstrating the validity of intervention-as-test is by using evidence-based interventions and by ensuring that, in each instance, they are implemented with fidelity. In this regard, the

standard treatment protocol may have a leg up on problem solving. Everyone knows what to implement because there is but one protocol, which makes training easier to accomplish and fidelity of implementation easier to assess and ensure; in turn, this makes it more likely that it can be "scaled up" in a district or school building. As best we know, however, the comparative fidelity of implementation (and effectiveness) of the two approaches has not been explored within the same experimental design. Such exploration represents an important and promising area of research.

Why RTI?

Special education costs

For decades, policymakers and academics have been frustrated by the LD construct generally and by IQ–achievement discrepancy particularly. One prominent reason is economics. In a sense, LD became too successful for its own good—if success may be defined by the number of children with the label. Shortly after LD was legitimized as a special-education category in the Education of All Handicapped Children Act of 1975, the proportion of children with LD in the general U.S. population skyrocketed from less than 2% in 1976–1977 to more than 6% in 1999–2000. This increase has proved expensive for school districts because, on average, it costs two to three times more to teach children with disabilities. Not long ago, New York City was spending US$1.67 billion, or 22 cents of every school dollar, on special education (Dillon, 1994) to provide services to 130,037 students or 13% of the city's one million school children (National Association of State Boards of Education, 1991).

IQ–achievement discrepancy

IQ–achievement discrepancy, which is the most widely used method of LD identification, has often been viewed as the culprit with respect to rising special education enrollments and costs, which brings us to a second reason for dissatisfaction with the LD construct. The discrepancy approach has been frequently criticized as atheoretical (e.g., Lyon, 1987; Willson, 1987) and, according to some, this absence of theory has permitted states and districts to specify discrepancy differently. Today, discrepancy varies nationwide in terms of (a) how it is computed (e.g., standard IQ score minus standard achievement score versus the regression of IQ on achievement), (b) its size (e.g., 1.0 *SD* versus 2.0 *SD*s), and (c) which IQ and achievement tests are used. Not surprisingly, these varying definitional features and criteria have led to large inconsistencies in LD prevalence between states and sometimes between districts within states (e.g., Reschly & Hosp, 2004; Scruggs & Mastropieri, 2002).

Such inconsistency in the definition of IQ–achievement discrepancy and varying prevalence rates—as well as the outright noncompliance by some school-level personnel with state and district guidelines (cf. Gottlieb, Alter, Gottlieb, & Wishner, 1994)—have contributed to a widespread view that the LD designation is whatever teachers and parents want it to be (e.g., Coles, 1987). A more damaging assertion, perhaps, is that the IQ–achievement discrepancy approach fails to distinguish a qualitatively different and more deserving subgroup of students from a much larger group of low achievers. Studies suggest that young, poor readers with and without an IQ–achievement discrepancy perform similarly on many reading-related cognitive tasks (e.g., Fletcher et al., 1994; Foorman, Francis, & Fletcher, 1995; Stanovich & Siegel, 1994), as well as demonstrate phonological processing deficits that are correctable with appropriate instruction (e.g., Fletcher, 1995; Morris et al., 1998; Stanovich, 1999; Torgesen, Morgan, & Davis, 1992; Vellutino et al., 1996).

Thus, say critics, thanks to the IQ–achievement discrepancy approach, the LD label is not just arbitrarily assigned, it is unfairly withheld from children who are as needy and deserving as those given the label. Many who advocate on behalf of RTI may view it as a means of reallocating resources—away from discrepant, middle class children of dubious disability to nondiscrepant, low-socioeconomic-status, low-achieving students who, prior to IDEA reauthorization, often fell between the cracks of service-delivery systems.

Concerns that IQ–achievement discrepancy is atheoretical and arbitrary, and that some of its basic assumptions have not been supported by research, have crystallized for many into two major criticisms. First, it represents a wait-to-fail model antithetical to early intervention; that is, children must fall dramatically behind their peers in academic achievement to qualify as LD. Second, critics say that the low achievement of so-called children with LD is presumed to reflect disability when, more times than not, it reflects poor teaching. Because RTI encourages appropriate use of evidence-based instruction across tiers, it should in principal decrease the numbers of children incorrectly identified as disabled.

Unanswered questions, unresolved issues

Problem solving and standard treatment protocol: Conceptual issues associated with "nonresponsiveness"

False negatives versus false positives

As described in a previous section, problem-solving and standard treatment protocol approaches differ operationally. They also differ conceptually, and each promotes a different implicit meaning of "nonresponsiveness." The standard treatment protocol may be considered a relatively rigorous test for nonresponsiveness and disability. Nonresponders, like those in the Vellutino et al. (1996) study, participated in evidence-based instruction delivered to small groups. The nonresponders' lack of progress almost seemed to defy the systematicity and intensity of their educational experience and the expertise and effort of their instructors. Their nonresponsiveness appears much more likely caused by disability than to the absence of good instruction. The standard treatment protocol seems to facilitate identification of "true positives," or children truly in need of special (e.g., individualized) services.

At the same time, use of the standard treatment protocol approach raises this question: Is it possible that some children who are nonresponsive to Tier 1 instruction, but who become responsive in a second or third tier, still have a disability and, once returned to their classroom for instruction (without the intensity and systematic instruction of the standard treatment protocol), will again demonstrate the same learning problems that first marked them as candidates for Tier 2? That is, whereas the standard treatment protocol approach is likely to identify true positives, it also appears likely to identify false negatives, or children who in higher tiers seem responsive and nondisabled but who, nevertheless, cannot survive in the mainstream classroom. The investigation by Vaughn et al. (2003) provides evidence in support of this possibility. A subset of children who met criteria for dismissal from intensive tutoring failed to perform adequately when they returned to their classrooms, and they eventually required additional tutoring.

Problem solving, with its typically *less* intensive and *less* systematic instruction, seems less likely than the standard treatment protocol approach to identify false negatives and more likely to identify false positives, or children who appear nonresponsive and disabled but, with more intensive instruction, can demonstrate they are neither, which raises the following question for practitioners and policymakers: Is it more desirable to err by identifying more false negatives (standard treatment protocol) or by identifying more false positives (problem solving)?

Who is in the normative population?

Problem-solving and standard treatment protocol approaches also pose different technical challenges. As indicated, the different problem-solving tiers typically occur in the classroom; standard treatment protocols are usually implemented outside the classroom in small groups. Assessing responsiveness to instruction in a classroom context has the advantage of a normative framework referenced to the larger population of typical students in school. That is, responsiveness to generally effective classroom instruction can be estimated for all students so that a normative profile can be generated to describe the full range of responses. With classroom instruction as the intervention, traditional cut points (e.g., 1.5 standard deviations below the mean) may be used to define disability. Such an approach requires measurement of all students. By contrast, it seems unlikely that a normative framework may be applied to the standard treatment protocol approach. Logistics and logic seem to argue against exposing the full range of students to an intensive tutoring regimen for the purpose of producing a normative profile. In all likelihood, practitioners would need to rely on a normative framework restricted to very poor readers, a proposition requiring empirical validation.

Access to special education

In comparison to the standard treatment protocol approach, problem solving usually represents a lower bar in determining nonresponsiveness and access to special education. Assuming that special education is effective, this helps ensure that all children with special needs receive appropriate services. Yet, relatively easy access to special education can, in some cases, reflect a rush to judgment and, as explained, the identification of false positives, or children who are incorrectly identified. In selecting between problem-solving and standard treatment protocol, it may be necessary to determine whether one's primary intent is identification or prevention.

Measuring and defining nonresponsiveness

Regardless of which RTI approach is adopted, two components of the assessment process must be specified. First, a method must be conceptualized for

measuring students' responsiveness to instruction. Second, once student responsiveness has been conceptualized, a criterion must be applied for defining nonresponsiveness. Beneath such a criterion, students are identified as having reading disabilities.

Various methods are available for specifying these two assessment components. As described in a previous section, Vellutino et al. (1996) tested students on WRMT–R several times over the course of a multiyear study. To establish a cut point for responsiveness, Vellutino et al. rank-ordered slopes representing children's growth in responsiveness to tutoring, performed a median split on the slopes, and designated the bottom half as nonresponsive. Similarly, Torgesen and colleagues (2001) evaluated student performance at the end of treatment on the subtests of the WRMTs, designating nonresponsiveness as failing to achieve so-called normalized status; that is, a word-reading standard score of 90 or better. Good, Simmons, and Kame'enui (2001), like Torgesen et al. (2001), also specified nonresponsiveness in terms of posttreatment status. However, their approach involves a criterion-referenced benchmark associated with future reading success.

Speece and Case (2001) took yet a different tack. They adopted frequent student monitoring using curriculum-based measures so that nonresponsiveness could be identified earlier in the school year than is possible with the Vellutino et al. (1996), Torgesen et al. (2001), or Good et al. (2001) methods. Speece and Case applied a "dual discrepancy" criterion (e.g., Fuchs & Fuchs, 1998; Fuchs, Fuchs, & Speece, 2002). Nonresponders were students whose slope and level of performance fell at least 1 standard deviation below their class mean. (This dual-discrepancy approach could also be determined with respect to school, district, or national norms or using benchmark cut points associated with future reading performance.)

These alternative methods produce different prevalence rates of reading disability and different subsets of nonresponsive children; that is, different children are identified by the different methods (see Fuchs, Fuchs, & Compton, 2004). This is important because a major criticism of IQ–achievement discrepancy as a method of LD identification has been the unreliability of the diagnosis. Practitioners relying on an assortment of assessment procedures in an RTI framework may produce similarly unreliable diagnoses. So researchers must develop a common approach to define and assess nonresponsiveness. Further complicating this task is that various assessment methods demonstrate differential utility in distinguishing responsive and nonresponsive groups on different components of reading; that is, an assessment method with demonstrated validity for beginning decoding skills may be invalid for assessing reading comprehension (see Fuchs et al., 2004). For this reason, consistency in identifying nonresponders across the various components of reading is an important criterion for selecting a valid approach to assessment.

These unanswered questions and issues challenge those who would use an RTI framework to define disabilities. Nevertheless, we believe the framework has strong potential. Right now, we most clearly see its promise in regards to how its multilayered structure can be implemented in the early grades to strengthen the intensity and effectiveness of reading instruction for at-risk students, preventing chronic school failure that corrodes children's spirit and diminishes all of us who work on behalf of the public schools.

DOUGLAS FUCHS is Nicholas Hobbs Chair of Special Education and Human Development at Vanderbilt University, where he codirects the Vanderbilt-Kennedy Center Reading Clinic. His research focuses on issues of disability classification and reading instruction for struggling learners. He can be contacted at the Department of Special Education, Box 328 Peabody, Vanderbilt University, Nashville, TN 37203, USA, or by e-mail at doug.fuchs@vanderbilt.edu.

LYNN S. FUCHS is Nicholas Hobbs Professor of Special Education and Human Development at Vanderbilt University. Her research focuses on the connection between classroom assessment and teachers' instructional planning as well as reading and mathematics disability. She can be contacted at 328 Peabody, Vanderbilt University, Nashville, TN 37203, USA, or by e-mail at lynn.fuchs@vanderbilt.edu.

REFERENCES

AL OTAIBA, S., & FUCHS, D. (in press). Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. *Journal of Learning Disabilities.*

COLES, G. (1987). *The learning mystique: A critical look at learning disabilities.* New York: Pantheon.

DILLON, D. (1994, August 14). New York City schools monitor contends that special education cost is inflated. *The New York Times*, p. 18.

DIVISION OF LEARNING DISABILITIES OF THE COUNCIL FOR EXCEPTIONAL CHILDREN. (n.d.). Response to OSEP's consensus statement. Unpublished paper.

EDUCATION OF ALL HANDICAPPED CHILDREN ACT OF 1975, Pub. L. 94–142 (5.6)

FLETCHER, J.M. (1995). *Diagnostic utility of intelligence testing and the discrepancy model for children with learning disabilities: Historical perspectives and current research.* Paper presented at the IQ Testing and Educational Decision Making Workshop, National Research Council, National Academy of Sciences, Washington, D.C.

FLETCHER, J.M., SHAYWITZ, S.E., SHANKWEILER, D.P., KATZ, L., LIBERMAN, I.Y., STUEBING, K.K., FRANCIS, D.J., ET AL. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology, 86*, 6–23.

FOORMAN, B.R., FRANCIS, D.J., & FLETCHER, J.M. (1995). *Growth of phonological processing skill in beginning reading: The lag versus deficit model revisited.* Paper presented at the Society for Research on Child Development. Indianapolis, IN.

FUCHS, D., FUCHS, L.S., & COMPTON, D.L. (2004). Identifying reading disability by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly, 27*, 216–227.

FUCHS, D., MOCK, D., MORGAN, P.L., & YOUNG, C.L. (2003). Responsiveness-to-intervention: Definitions, evidence, and

implications for the learning disabilities construct. *Learning Disabilities Research & Practice*, *18*, 157–171.

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

FUCHS, L.S., FUCHS, D., & SPEECE, D.L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly*, *25*, 33–46.

GOOD, R.H., SIMMONS, D.C., & KAME'ENUI, E.J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*, 257–288.

GOTTLIEB, J., ALTER, M., GOTTLIEB, B.W., & WISHNER, L. (1994). Special education in urban America: It's not justifiable for many. *Journal of Special Education*, *27*, 453–465.

GRIMES, J. (2002). *Responsiveness to interventions: The next step in special education identification, service and exiting decision making*. Paper presented at Office of Special Education Programs LD Summit Conference, Washington, DC.

IKEDA, M., & GUSTAFSON, J.K. (2002). *Heartland AEA 11's problem solving process: Impact on issues related to special education* (Research Rep. No. 2002–01). Johnston, IA: Heartland Area Education Agency 11.

INDIVIDUALS WITH DISABILITIES EDUCATION IMPROVEMENT ACT OF 2004, Pub. L. 108–466.

LYON, G.R. (1987). Severe discrepancy: Theoretical, psychometric, developmental, and educational issues. *Learning Disabilities Research & Practice, 3*, 10–11.

LYON, G.R. (1995). Research initiatives in learning disabilities: Contributions from scientists supported by the National Institute of Child Health and Human Development. *Journal of Child Neurology*, *10* (suppl. 1), S120–S126.

MARSTON, D., MUYSKENS, P., LAU, M., & CANTER, A. (2003). Problem-solving model for decision making and high-incidence disabilities: The Minneapolis experience. *Learning Disabilities Research & Practice, 18*, 187–200.

MCMASTER, K.L., FUCHS, D., FUCHS, L.S., & COMPTON, D.L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children*, *71*, 445–463.

MORRIS, R.D., SHAYWITZ, S.E., SHANKWEILER, D.R, KATZ, L., STUEBING, K.K., FLETCHER, J.M., LYON, R.G., ET AL. (1998). Subtypes of reading disability: Variability around a phonological core. *Journal of Educational Psychology*, *90*, 347–373.

NATIONAL ASSOCIATION OF STATE BOARDS OF EDUCATION. (1991, October). *Special education: New questions in an era of reform*. Alexandria, VA: Author.

NO CHILD LEFT BEHIND ACT OF 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).

O'CONNOR, R.E. (2000). Increasing the intensity of intervention in kindergarten and first grade. *Learning Disabilities Research & Practice*, *15*, 43–54.

RESCHLY, D.J., & HOSP, J.L. (2004). State SLD identification policies and practices. *Learning Disability Quarterly*, *27*, 197–213.

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

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.

STANOVICH, K.E. (1999). The sociopsychometrics of learning disabilities. *Journal of Learning Disabilities*, *32*, 350–361.

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.

TELZROW, C.F, MCNAMARA, K., & HOLLINGER, C.L. (2000). Fidelity of problem-solving implementation and relationship to student performance. *School Psychology Review*, *29*, 443–461.

TORGESEN, J.K., ALEXANDER, A., WAGNER, R., RASHOTTE, C., VOELLER, K., & CONWAY, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities*, *34*, 33–58.

TORGESEN, J.K., MORGAN, S., & DAVIS, C. (1992). The effects of two types of phonological awareness training on word learning in kindergarten children. *Journal of Educational Psychology*, *84*, 364–370.

VAUGHN, S., LINAN-THOMPSON, S., & HICKMAN, P. (2003). Response to instruction as a means of identifying students with reading/learning disabilities. *Exceptional Children*, *69*, 391–409.

VELLUTINO, F.R., SCANLON, D.M., SIPAY, E.R., SMALL, S., 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.

WILLSON, V.L. (1987). Statistical and psychometric issues surrounding severe discrepancy. *Learning Disabilities Research*, *3*(1), 24–28.