Learning Disabilities Research & Practice, *18*(3), 157–171

Responsiveness-to-Intervention: Definitions, Evidence, and Implications for the Learning Disabilities Construct

Douglas Fuchs
George Peabody College of Vanderbilt University

Devery Mock
University of Virginia

Paul L. Morgan and Caresa L. Young
George Peabody College of Vanderbilt University

Abstract. Longstanding concern about how learning disabilities (LD) are defined and identified, coupled with recent efforts in Washington, DC to eliminate IQ-achievement discrepancy as an LD marker, have led to serious public discussion about alternative identification methods. The most popular of the alternatives is responsiveness-to-intervention (RTI), of which there are two basic versions: the "problem-solving" model and the "standard-protocol" approach. The authors describe both types, review empirical evidence bearing on their effectiveness and feasibility, and conclude that more needs to be understood before RTI may be viewed as a valid means of identifying students with LD.

THE RISE AND (APPARENT) FALL OF IQ-ACHIEVEMENT DISCREPANCY

The Rise . . .

In the mid-1970s, advocates convinced Congress to include learning disabilities (LD) in the new Education for All Handicapped Children Act (PL 94-142) by arguing that LD represented a unique group of children: those demonstrating unexpected learning failure and specific learning failure (e.g., Kavale, 1987). There were two compelling reasons, the advocates claimed, to view "unexpected" and "specific" learning failure as signature features of the disability. First, as far back as the 1890s, physicians W. Pringle Morgan and John Hinshelwood each described the seemingly paradoxical inability of some children of average and superior intelligence to master academic concepts, a phenomenon documented more extensively by another physician, Samuel Orton, in the 1920s and 1930s (Hallahan & Mercer, 2001).

Second, in 1975, Rutter and Yule reported findings from a series of epidemiological studies that seemed to buttress Morgan, Hinshelwood, and Orton's clinical observations. Rutter and Yule measured the IQ and reading performance of all 9 and 14 year olds on the Isle of Wight. The researchers regressed the children's IQ scores on their reading scores to produce a distribution of IQ-predicted reading performance. Scores above the mean represented overachievement (i.e., exceeding prediction); scores below the mean indicated underachievement (i.e., beneath prediction). Whereas such a distribution should resemble a Gaussian curve, with overachievement occurring as frequently as underachievement, Rutter and Yule reported a "hump" at the lower end of the distribution, which, they said, indicated that "extreme degrees" of reading underachievement occur at a greater rate than should be expected (Rutter & Yule, 1975, p. 185). When Rutter and Yule compared the "underachievers" to the children whose low reading performance was commensurate with their equally low IQ scores (i.e., "low achievers"), they found that the underachievers were different "in terms of sex distribution, neurological disorder, and pattern of neuro-developmental deficit" (p. 194). Further, the underachieving readers had a worse prognosis for reading and spelling and a better prognosis for mathematics. This led Rutter and Yule to suggest that the group of underachievers, or children with "specific reading retardation," was distinctly different from the group of low achievers, or "generally backward readers." Findings appeared to confirm "unexpected" and "specific"

Requests for reprints should be sent to Douglas Fuchs, 328 Peabody, Vanderbilt University, Nashville, TN 37203. Electronic inquires may be sent to doug.fuchs@vanderbilt.edu.

learning failure as valid identifiers, or markers, of students with LD (e.g., Fletcher, 1995).

In 1977, the U.S. Office of Education proposed regulations to accompany PL 94-142 to help practitioners identify children with LD. The government's strategy, no doubt influenced by Rutter and Yule's (1975) research, was to suggest a "severe discrepancy" between performance on intelligence and achievement tests as the primary marker of LD. Government officials wrote more specifically that educators may identify children as LD if the children receive appropriate learning experiences for their age and ability and still do not achieve commensurate with their age and ability in oral expression, listening comprehension, written expression, basic reading skills, reading comprehension, mathematics education, or mathematics reasoning.

Most state departments of education adopted the severe discrepancy idea (Frankenberger & Franzaglio, 1991; Mercer, King-Sears, & Mercer, 1990), but they defined it in their own ways. Their definitions of discrepancy varied in terms of how it was computed (e.g., standard IQ score minus standard achievement score vs. the regression of IQ on achievement), its size (e.g., 1.0 *SD* vs. 2.0 *SD*s), and which specific IQ and achievement tests were used. Not surprisingly, the application of these varying definitional features and criteria led to large inconsistencies in LD prevalence between states (e.g., Scruggs & Mastropieri, 2002).

And Fall

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; MacMillan, Gresham, & Bocian, 1998; Shepard & Smith, 1983)—have contributed to a widespread view that the LD designation is arbitrary (e.g., Coles, 1987; Reynolds, Wang, & Walberg, 1987; Skrtic, 1991; Sleeter, 1998; Ysseldyke, Algozzine, & Epps, 1983). A more damning 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—Rutter and Yule's (1975) research notwithstanding. Recent 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; Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Stanovich & Siegel, 1994), as well as demonstrate phonological-processing deficits, which are correctable with appropriate instruction (e.g., Fletcher, 1995; Morris et al., 1998; Siegel, 1989, 1999; Stanovich, 1999; Torgesen, Morgan, & Davis, 1992; Vellutino et al., 1996).

Thus, say the 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 to whom the label is given. Many of the deserving but unidentified students are from low-income homes with relatively low IQ scores insufficiently different from their low achievement scores to qualify them for special education services. Commenting on this apparent class bias, Stanovich (1999) writes, "[I]t is rare for the advocates of discrepancy-based definitions to articulate the theory of social justice that dictates that society has a special obligation to bring up the achievement of individuals whose achievements fall short of their IQs, rather than simply to bring up the skills of those with low skills, period" (p. 353).

More concerns have been expressed. Critics charge that IQ tests are a poor index of intelligence; the IQ-achievement discrepancy approach represents a "wait-to-fail" model since many students must perform poorly for years before their achievement scores are sufficiently below their IQ scores; the low achievement of many children with the label reflects poor teaching rather than a disability, despite federal regulations requiring that appropriate learning experiences be a precondition of labeling; and much of the necessary and costly data collection has as little to do with instruction as the label itself. That is, many ask: What set of special and effective practices are exclusively associated with "LD"?

Dissatisfaction is coming to a head for at least two reasons. First, throughout the 1980s and 1990s, increasing numbers of students with the LD label and escalating special education costs vexed local and state educational and political leaders. IQ-achievement discrepancy—the most widely used method of LD identification—has often been viewed as the culprit. As importantly, educational advisors and appointees in the current administration in Washington, DC share many of the just-mentioned concerns about IQ-achievement discrepancy. They seem determined to reduce or eliminate its use by practitioners. Wade Horn, a member of the President's Commission on Excellence in Special Education, was quoted in the Commission's report as saying, "I would like to encourage the Commission to drive a stake through the heart of this overreliance on the discrepancy model for determining the kinds of children that need services" (p. 25; also see a widely circulated speech by the Assistant Secretary of Education, Robert Pasternack, entitled, "The Demise of IQ Testing for Children with Learning Disabilities").

Although some professional organizations and academics continue to support the IQ-achievement-discrepancy approach, most acknowledge it to be a flawed criterion. Evidence of this abounds. Speece and Shekitka (2002), for example, recently asked 218 members of the editorial boards of several scholarly journals how reading disabilities should be operationalized. Seventy percent claimed IQ-achievement discrepancy should play no role. Additionally, in Fall of 2001, the Office of Special Education Programs convened a group of researchers from across the country to establish parameters for discussions about LD as part of the

upcoming reauthorization of IDEA. A majority of the researchers rejected IQ-achievement discrepancy as a valid LD marker, a stance subsequently supported, more or less, by the National Joint Committee on Learning Disabilities, a group of major professional and advocacy organizations (see Bradley, Danielson, & Hallahan, 2002).

Responsiveness-to-Intervention

If we are to abandon a method of LD identification that has been used by hundreds of thousands of practitioners for more than a quarter-century, the obvious and critical questions become: What's the alternative? and How good is it? Responsiveness to intervention (RTI) is the front-running alternative. It has recently been discussed or advocated by the Division for Learning Disabilities of the Council for Exceptional Children (2002); the International Dyslexia Association (Dickman, Hennessy, Moats, Rooney, & Tomey, 2002); the National Association of School Psychologists (2002); the National Association of State Directors of Special Education (Corpolongo, Pochowski, Lloyd-Jones, & Lenz, 2002); the National Center for Learning Disabilities (2002); the National Research Council (Donovan & Cross, 2002); the Office of Special Education Programs in the U.S. Department of Education (2002; Bradley et al., 2002); and the President's Commission on Excellence in Special Education (2002). It has also been the focus of concept papers (e.g., L. S. Fuchs, 1995; L. S. Fuchs, Fuchs, & Speece, 2002) and recent research (e.g., Al Otaiba & Fuchs, 2003; McMaster, Fuchs, Fuchs, & Compton, 2003; Speece & Case, 2001; Speece, Case, & Molloy, in press; Vaughn & Linan-Thompson, 2003; Vellutino et al., 1996).

In broad terms, RTI may be described as follows.

1. Students are provided with "generally effective" instruction by their classroom teacher;
2. Their progress is monitored;
3. Those who do not respond get something else, or something more, from their teacher or someone else;
4. Again, their progress is monitored; and
5. Those who still do not respond either qualify for special education or for special education evaluation.

At present, RTI is implemented in various ways in at least a score of districts (Canter, 2003). The RTIs differ in terms of the number of levels in the process; who delivers the interventions; and whether the process is viewed as a precursor to a formal evaluation for eligibility, or if RTI is itself the eligibility evaluation.

According to supporters, RTI solves many problems associated with IQ-achievement discrepancy. Foremost, it is described as providing help more quickly to a greater number of struggling students. Related to this is the idea that by providing individualized and intensive instruction to these students, the approach effectively separates students with disabilities from those who perform poorly because of inadequate prior instruction. Third, this distinction between truly disabled children and "false positives" (children who appear to be disabled but are not), leads to a reduction in special education enrollment and cost. Fourth, the provision of services is not contingent on one's performance on an IQ test. And, finally, because some RTI approaches are noncategorical, they circumvent the use of putatively stigmatizing labels such as "LD."

At present, two identifiable groups have vigorously promoted RTI: an early intervention/prevention group consisting of early reading researchers and behaviorally-oriented school psychologists who have lobbied for years to change the psychometrically-driven identification process and their typical role as psychometrists. Whereas the early interventionists and school psychologists agree that the LD identification process needs fundamental revision, beginning with the elimination of IQ-achievement discrepancy, they have different visions of how RTI should play out. For the school psychologists, RTI is synonymous with "the problem-solving model"; for the early reading group, RTI is accomplished by use of a standard and validated-treatment protocol. As these groups advocate their respective strategies, they have inadvertently contributed to the confusion surrounding RTI.

Because the problem-solving model is the more widely used of the two approaches, the primary purpose of this article is to provide a detailed explanation of the problem-solving version and to review the evidence of its effectiveness and feasibility. We focus on four programs—in Ohio, Pennsylvania, Iowa, and Minnesota—that have been highlighted as exemplars of the problem-solving approach (cf. Horowitz, Lichtenstein, Roller, Toombs, & Kukic, 2002). Although our take on the evaluations of, and evidence for, these programs is sometimes critical, we wish readers to know that we hold all the practitioners who developed them in high regard. Their efforts reflect a serious and noble commitment to help children, and especially low-performing children, achieve at much higher levels. After discussing the problem-solving programs, we describe more briefly the standard-protocol approach, and we close by drawing implications of RTI for the LD construct.

PROBLEM SOLVING

The Model

According to Telzrow, McNamara, and Hollinger (2000), problem solving was first described in the behavioral consultation literature. As explained by Bergan (1970), Bergan and Kratochwill (1990), Tharp and Wetzel (1969), and others, behavioral consultation's problem solving is inductive, empirical, and, of course,

behavioral. But the key feature here is *inductive*. Supporters of the model believe no student characteristic (e.g., disability label, race, SES, neighborhood) dictates *a priori* what intervention will work. Nor will a given intervention be effective for all students of a particular group, irrespective of how exclusively the group may be conceived. Instead, solutions to instructional and behavioral problems are induced by evaluating students' responsiveness to a four-stage process comprising problem identification, problem analysis, plan implementation, and problem evaluation. The major objectives of *problem identification* are to define the problem behavior in observable terms and obtain a reliable estimate of its frequency, intensity, or duration. In the *problem-analysis* stage, the goal is to validate the existence of the problem, identify instructional and student variables that may contribute to a solution, and develop an appropriate plan. During *plan implementation*, the consultant monitors implementation and provides corrective feedback, helping ensure that the intervention is delivered as designed. And in *problem evaluation*, the consultant and teacher evaluate the effectiveness of the intervention and, if it has proved ineffective, determine how it should be modified. In short, in the behavioral problem-solving approach, there may be numerous potential solutions to a given problem. Successful solutions are often achieved through trial and error, which is not to suggest random, or unthinking, approaches to instruction.

These trial-and-error approaches rely on the careful collection of data on students' deportment or academic performance in response to treatment. To establish a baseline, for example, practitioners need evidence like "Dolly was off task 75 percent of the time she was observed during math instruction; her peers, only 25 percent." Or, "She correctly reads three words per minute in first-grade text, whereas the local norm is 27 words." These comparisons facilitate both judgments about the seriousness of the problem and goal setting (e.g., "Dolly's 50 percent discrepancy will be reduced by half in three weeks"; or, "The discrepancy of 24 words read correctly will be reduced to 12 words in three weeks of peer tutoring"). It also illustrates the "behavioral" in behavioral problem solving: defining problems in the context of person-environment situations rather than attributing them to putatively fixed student characteristics like mental retardation, insolence, or a "bad" family.

Finally, behavioral problem solving is "triadic," involving a consultant, teacher, and student. The consultant's effect on the student is indirect, that is, the teacher mediates between the consultant's contribution and the student. And the consultant-teacher relationship is supposed to be collegial. The consultant generally has no formal authority over the teacher, and the teacher's participation is voluntary (Bergan & Kratochwill, 1990, p. 14).

A sizable research literature explores the effectiveness of behavioral problem solving. Some of these studies have been judged stronger in design than others (for critiques of the literature, see Alpert & Yammer, 1983; Erchul & Martens, 2002; Meyers, Pitt, Gaughan, & Friedman, 1978; Sheridan, Welch, & Scott, 1996; West & Idol, 1987; Witt, Gresham, & Noell, 1996). Nevertheless, quantitative (e.g., Medway & Updyke, 1985) and qualitative (e.g., Sheridan et al., 1996) reviews of these studies come to the same important conclusion: consultants using behavioral problem solving are frequently successful in addressing a wide range of student problems, with the result that teachers and consultants regard it as worthwhile (e.g., Kratochwill, Elliott, & Rotto, 1995; Kratochwill & Van Somersen, 1985).

Problem Solving and Prereferral Intervention

Between 1977 and 1994, students with disabilities increased from 3.7 million to 5.3 million "despite . . . that overall public school enrollment [remained] roughly constant over this period" (Hanushek, Kain, & Rivkin, 2001, p. 7). This represented an increase from 8.3 percent to 12.2 percent of the general student population. Virtually all the growth came from increases in students classified as LD, a group that grew from 22 percent to 46 percent of all special-needs children over this period (Hanushek et al., 2001, p. 7). Hanushek and Rivkin (1997, cited in Hanushek et al., 2001) suggested that "special education accounted for roughly 20% of the increase in per student spending during the 1980s, slightly less than double the share of special education students" (p. 7). These developments did not escape the attention of school boards, school superintendents, politicians, and other stakeholders in public education, some of whom began calling for the immediate downsizing of special education (see Viadero, 1991).

By the mid-1980s, prereferral intervention was seen by policymakers, academics, and practitioners as a solution to the apparent overidentification of students with disabilities. "Prereferral intervention" refers to a teacher's modification of instruction, or some other aspect of the learning environment, to better accommodate a difficult-to-teach student prior to a formal referral of the student for testing and possible special education placement. This prereferral activity is often "brokered" by a consultant or team working indirectly with a targeted difficult-to-teach student through consultation with the teacher. Implicit is a preventive intent, that is, eliminating inappropriate referrals while increasing the legitimacy of those that are initiated and reducing future student problems by strengthening the teacher's capacity to intervene effectively with a greater range of children.

Researchers, many of whom were supported by the Office of Special Education Programs (cf. Kaufman, Kameenui, Birman, & Danielson, 1990), developed two versions of prereferral intervention. One followed the dictates of behavioral consultation's problem solving (e.g., D. Fuchs & Fuchs, 1989; D. Fuchs, Fuchs, & Bahr, 1990; D. Fuchs, Fuchs, Bahr, Fernstrom, &

Stecker, 1990; D. Fuchs, Fuchs, Harris, & Roberts, 1996; Graden, Casey, & Bonstrom, 1985; Ponti, Zins, & Graden, 1988); the other reflected "collaborative consultation" (e.g., Friend & Cook, 1992; Pugach & Johnson, 1995), which gives more attention to interpersonal relations. Perhaps the best-known example of collaborative consultation is Chalfant and colleagues' Teacher Assistance Teams (e.g., Chalfant, Pysh, & Moultrie, 1979). As the name suggests, it is a team approach meant to be "*of* teachers and *for* teachers." Specialists typically do not participate. Collaborative consultation generally, and Teacher Assistance Teams in particular, have been infrequently researched (cf. Sheridan et al., 1996).

Collaborative Problem Solving

Rather than choose between these alternatives, some states and districts have developed a hybrid of behavioral problem solving and Teacher Assistance Teams, calling their creation "collaborative problem solving." The admixture is illustrated by the use of multidisciplinary teams, composed of specialists (behavioral problem solving) and teachers (collaborative consultation) who are trained in both the four-stage problem-solving process (behavioral problem solving) and interpersonal relations (collaborative consultation). These multidisciplinary teams have become popular in part because they are a relatively efficient means of delivering prereferral intervention to many teachers in a school and to many schools in a district. They also reflect the spirit of the ongoing reform movement in education; specifically, notions of collegiality, bottom-up decision making, and egalitarianism (see D. Fuchs & Fuchs, 1996). Kovaleski (2002) expresses this perspective when writing about the importance of teams: "Facilitating change in general education classrooms can be daunting work, and the team structure allows for members to support each other in the face of unavoidable resistance to change" (p. 648). "Teams," he writes, "[allow] for the development of a sense of mission and team spirit" (p. 648).

Despite their popularity, or maybe because of it, important questions should be asked of these teams. In the behavioral problem-solving literature, a consultant and teacher typically work through a four-step, inductive, data-based process. Do these important features or other aspects of the process change significantly in the context of a multidisciplinary team? In the research on behavioral problem solving, teacher participation is always voluntary. It is not voluntary in many schools and districts where, if a teacher wants to refer Johnny, she or he must agree to work with a team to modify classroom curricula, materials, or instruction. Does the involuntary nature of this relationship affect teachers' treatment implementations? What child information does the team collect to explore responsiveness to intervention? With these and other questions in mind, we turn to Ohio and Pennsylvania, two states in which practitioners have implemented collaborative problem solving on a large scale and have tried hard to evaluate its effects.

Ohio's Intervention Based Assessment (IBA)

IBA began in 1992–1993 as a voluntary school-based initiative under a special education waiver plan by the Ohio State Department of Education. Its purpose was to produce treatment plans for nondisabled students with behavior or learning problems, or to be used as part of a multifactored evaluation for children with suspected disabilities to identify effective interventions to be incorporated into their IEPs (Telzrow et al., 2000, p. 445). Schools were invited to participate in the state's initiative if they already had problem-solving teams in place. In 1992–1993, a first cohort of 35 schools was recruited statewide.

According to Telzrow et al. (2000), IBA combines a behavioral problem-solving approach with collaborative consultation. Its problem-solving components include: a behavioral definition of the problem; baseline data; explicit goal setting; an hypothesized reason for the problem; an intervention plan; evidence of fidelity of treatment implementation; data indicating student responsiveness to treatment; and comparison of student performance to baseline. These activities are conducted by multidisciplinary teams that, at minimum, include the principal, school psychologist, special education teacher, and classroom teacher.

In 1996–1997, Telzrow et al. (2000) conducted a statewide evaluation of the IBA program. Among 329 identified IBA schools, 227 (69 percent) were selected for study. The remaining IBA schools (31 percent) were judged unready for evaluation. Study schools "were directed to submit 'best case' documentation (i.e., products that would reflect their most complete and accurate implementation of the problem-solving process), and had sole discretion over the selection of cases submitted" (Telzrow et al., 2000, p. 449). The schools had two instruments to help them with their documentation. These instruments directed the schools to list each of their problem-solving components and to describe their concerns about the student's learning, the chosen intervention, how its implementation was monitored, and its effectiveness. A five-point Likert scale and scoring rubric were developed to evaluate the fidelity of problem-solving implementation and the degree of student change during the IBA process. As indicated in Telzrow et al.'s Tables 2 and 3 (pp. 450–451), "Ohio's [multidisciplinary team's problem-solving implementation] was frequently inconsistent and below desired levels of fidelity" (p. 457). For example, Telzrow and her colleagues reported an average rating of 2.6 (out of a possible 5) for the problem-solving component requiring "evidence that the intervention was implemented as designed." This rating is worse than a 3.0 rating, which, according to the rubric used, would have indicated that the typical IBA documented treatment integrity "with

vague general statements" (p. 451) such as "with [the mere] assertion that the treatment occurred" (p. 451).

Telzrow and her colleagues concluded, "The present study suggests that reliable implementation of problem-solving approaches in schools remains elusive" (p. 458). And, "[a]lthough problem-solving consultation has considerable intuitive appeal, attributions of positive outcomes to such processes are not defensible until research confirms reliable and consistent implementation" (p. 457). What makes these findings all the more noteworthy is that Telzrow et al. deliberatively (and openly) stacked the deck in favor of the schools. School personnel were told to present their best-case documentation. Additionally, as expressed by Telzrow and associates, "[T]he fact that . . . [only] two thirds of all participating buildings submitted case documentation may suggest that the sample used in this analysis reflects a bias towards more proficient [multidisciplinary teams]" (p. 453).

Pennsylvania's Instructional Support Teams (ISTs)

Begun in 1990 by Jim Tucker, Director of the Bureau of Special Education in the Pennsylvania Department of Education (Conway & Kovaleski, 1998), the ISTs are maybe the best-known statewide prereferral intervention program in the nation. Districts in other states (e.g., New York; see Papandrea, Walkley, DeLorme, & McNamara, 2001) have used IST as a model for their own efforts. Like Ohio's IBAs, the ISTs are an example of collaborative problem solving in the service of providing prereferral intervention. Unlike Ohio's IBAs, ISTs include a support teacher whose primary responsibility is to help the classroom teacher implement the prereferral intervention. In addition to the support teacher, ISTs comprise the student's teacher, a principal, and specialists as needed (Kovaleski, Gickling, Morrow, & Swank, 1999). All are trained in problem solving, communication skills, and collaboration (Conway & Kovaleski, 1998).

The ISTs use curriculum-based assessment (for academic concerns) and behavioral assessment (for behavior concerns) to describe a student's problem in measurable terms. A goal is then set and, based on the assessment data, an intervention plan is developed. The classroom teacher and support teacher work with the student in the classroom, and the teachers continuously monitor the child's progress to help the team determine intervention effectiveness. This instructional support is limited to 50 school days, after which the IST meets to review the student's progress and determine whether further evaluation is necessary. Students making little progress are given a multidisciplinary evaluation for possible special education placement (Conway & Kovaleski, 1998, p. 347).

Kovaleski, Tucker, and Duffy (1995, cited in Conway & Kovaleski, 1998) report very high fidelity of implementation among IST members across the state. "In the second year of implementation," Kovaleski et al. write, "teams of training consultants and members of ISTs from other areas of the state visited the schools . . . . These teams assessed essential program elements through direct observation, interviews, and record interviews. Over the course of the project, 98 percent of the schools achieved validation on their first visit" (p. 349). We could not find an explanation of what schools had to do to achieve "validation" or whether and how the observation and interview data were combined to produce an index of implementation fidelity. Nor did we find actual treatment integrity data. In the absence of this information, we do not really know what to make of Kovaleski et al.'s finding that "98 percent of the schools achieved validation on their first visit."

Hartman and Fay (1996) studied IST implementation in 1,074 schools in 1992–1993 and 1993–1994 (p. 7). They reported that the use of ISTs led to fewer special education referrals, a decrease in special education placements, and a reduction in grade retentions. Reflecting on these results, Kovaleski et al. (1999) write, "they represent only indirect reflections of the effectiveness of the IST process" (p. 172). "[Such] data," they continued, "can be interpreted as beneficial only if there is clear evidence that students not referred or retained are successful in general education programs" (p. 172). To explore student success, Kovaleski et al. studied the impact of IST interventions on students' time on task, task completion, and task comprehension. Students in schools in which ISTs were implemented with high fidelity showed better progress on these measures than did students in non-IST schools. We could not find information on either the nature of the treatments implemented or data on interrater agreement in connection with the implementation data. More importantly, time on task, task completion, and task comprehension are only indirect academic measures. With that said, in comparison to most district and state evaluations of prereferral intervention programs, Kovaleski et al.'s findings are among the more persuasive.

Summary

A prevalent view in the 1980s was that too many children were referred for testing and possible special education placement. This apparent overreferring was interpreted as a sign of many teachers' inability or unwillingness to accommodate academic diversity in their classrooms (cf. Reynolds et al., 1987). Federal dollars encouraged (mostly special education) researchers to develop prereferral intervention programs and, by the mid-1980s, two types of programs emerged: behavior problem solving and collaborative consultation. Both versions were conceived as means to increase teacher support prior to formal referral making and in both students unresponsive to classroom interventions were seen as appropriate candidates for testing and possible special education placement. Many practitioners began implementing a hybrid of the two

types of prereferral intervention, calling it collaborative problem solving. Signs are that collaborative problem solving is popular among many practitioners. However, practitioners (and researchers) have infrequently undertaken evaluations of this approach. The few who have done so have generally failed to produce persuasive evidence that classroom-based interventions (1) are implemented with fidelity and (2) strengthen students' academic achievement or improve classroom behavior. Our purpose, of course, is not to belittle such efforts. Conducting large-scale implementations and evaluations of prereferral intervention requires considerable resources, and it is the rare school district that has them. Those who have researched their respective programs have tended to demonstrate remarkable perseverance and professionalism and should be commended. But none of this diminishes the fact that, as we write, many practitioners are using unvalidated prereferral intervention processes. More troubling, and more to the point of this article, is that these largely untested procedures are the basis of a much more ambitious and complex form of RTI, which is currently being considered as a replacement for IQ-achievement discrepancy to identify students with LD.

Problem Solving as Eligibility Decision Making

Grimes (2002) writes that a small but growing number of school districts and educational agencies in Florida, Iowa, Kansas, Minnesota, South Carolina, and Wisconsin are using three- or four-level versions of Ohio's and Pennsylvania's collaborative problem solving to both provide support to classroom teachers and to identify students for special education services. In accordance with these more ambitious and complex models, when a child fails to respond at one level of assistance, more resources and greater expertise are brought to bear at the next higher level. "As the magnitude of students' needs increases, the amount of educational resources increases proportionally to provide interventions required to support students' . . . growth" (Grimes, 2002, p. 8). In a few places, this extended problem-solving approach has become more than prereferral intervention; rather it has become prereferral intervention and the eligibility process all rolled into one. Best-known exemplars of this approach have been developed by the Heartland Area Educational Agency (hereafter, "Heartland") and Minneapolis Public Schools.

Heartland

Heartland, the largest of Iowa's 15 educational agencies (Ikeda, Tilly, Stumme, Volmer, & Allison, 1996), is responsible for providing training and technical assistance to 56 districts and 350 schools (Ikeda & Gustafson, 2002a). Statewide reform began in 1985 and has continued in earnest since 1992 (Ikeda & Gustafson, 2002a). As part of reform making, 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, at Level 1, a teacher confers with a student's parent(s) to try to resolve academic or behavior problems. At Level 2, the teacher and his or her school's Building Assistance Team (BAT) 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 are mostly doctoral-level or masters-level school psychologists and special educators who use behavioral problem solving to refine or redesign the intervention and coordinate its implementation from that point on. (Several of the first author's closest colleagues, who have frequently consulted with many of these practitioners, attest to this group's uncommon commitment, energy, organization, and professional savvy and skill.) Finally, at Level 4, special education assistance and due process protections are considered.

At each level, the problem-solving 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 based on student responsiveness; and evaluate its effectiveness and plot future actions (cf. Grimes, 2002). Throughout this problem-solving process, and across the four levels, "data about a student's responsiveness to intervention becomes the driving force" (Grimes, 2002, p. 4). Teachers and Heartland staff are directed to compare the student's performance level and learning rate with what is expected of students in the same classroom. It is the student's relative classroom performance, rather than test performance, that determines responsiveness/unresponsiveness and, eventually, special education eligibility. In other words, Heartland's multilevel approach is meant to provide teachers with repeated opportunity and increasing levels of support to help students become responsive to instruction. Students who remain unresponsive are eligible for special education.

According to Ikeda and Gustafson (2002b), Heartland's recursive, multilevel problem-solving model has several advantages over the traditional psychometric approach to eligibility decision making. First, the model directs practitioners to collect more thorough and instructionally relevant assessment data. Because cut scores (e.g., a standard score of 80 on an achievement test) have been deliberately excluded from consideration, the problem-solving model "force[s] decisions to be made using *multiple sources* that show a *convergence of evidence*" (Ikeda & Gustafson, p. 6, emphasis in original; also see Ikeda et al., 1996). Second, use of the model leads to a match between the presenting problem, assessment methods, and interventions. Third, the model is noncategorical. Children found eligible for special education services are not labeled LD, mentally retarded,

or behavior disordered. This facilitates the provision of services to at-risk students who might otherwise not receive them. Heartland staff would probably say that they are as interested in documenting "unexpected underachievement" as are supporters of traditional definitions of LD. However, as indicated, Heartland staff defines unexpected underachievement as a comparison between Tom's progress and his classmates' progress, rather than as a significant discrepancy between Tom's performance on IQ and achievement tests. In accordance with the Heartland model, IQ tests are virtually never administered.

Although Heartland has implemented versions of a multilevel problem-solving approach since 1985, "results of those efforts are largely undocumented" (Ikeda & Gustafson, 2002b, p. 1). Thus, Ikeda and Gustafson and their Heartland colleagues conducted evaluations in 1999–2000 and 2000–2001 "to present . . . initial data on the effects of this service delivery system on student outcomes and referral rates" (Ikeda & Gustafson, 2002b, p. 1). In Year 1 (1999–2000), Ikeda and Gustafson requested 10 percent of the BATs in Heartland to participate "based on recommendations from agency staff" (p. 9). Reasons for these recommendations were not given. Thirteen of 56 districts (23 percent) and 15 of 350 schools (4 percent) participated. Across the study schools, data were collected on as many as 344 students with BAT involvement. Ikeda and Gustafson write that for 19 percent of the students, problems were "resolved" (p. 11); for 39 percent, interventions were continuing; 4 percent were pending Level 3 assessment; and 25 percent were staffed into special education at the end of the year. Of the 344 students, 25 percent had previously been referred (p. 10).

In Year 2 (2000–2001), one elementary school in each of Heartland's 56 districts was sent a survey on the BATs. Twenty-nine schools in 28 districts agreed to participate. Respondents indicated that the BATs "resolved" 20 percent of their cases (Ikeda & Gustafson, 2002b, p. 14). An additional 41 percent of the students were involved in ongoing interventions; 8 percent were awaiting Level 3 assessments; and 22 percent had been staffed into special education (p. 14). Thirty-four percent of students had previously been referred (p. 13).

Summarizing the findings from their evaluations, Ikeda and Gustafson (2002b) write that, "The purpose of problem solving is to help . . . develop interventions to address problems in general education, leaving special education resources for the most severe problems. Two years of data . . . indicate that the number of problems addressed without special education resources is about 75%. [A] substantial number of children [are] being helped by area building assessment teams. [I]nterventions are being implemented with success in the general education classroom" (p.14). This conclusion notwithstanding, several important questions can be raised about the database: the authors do not mention the small number of districts, schools, teachers, and students who participated in the studies (e.g., only 4 percent of Heartland's schools were involved in the Year 1 evaluation); nor do they discuss whether the participating districts, schools, teachers, and students are representative of Heartland. Additionally, do all districts and schools implement the problem-solving model? Does each school conduct it pretty much the same way? We suspect not, based on this from Ikeda et al. (1996): "There is significant variability between individual school districts, and even in buildings within districts, in internalization of assumptions needed for using problem solving consultation" (p. 241).

In a different vein, we could not find any explanation of the authors' claim that BATs successfully addressed 75 percent of the problems brought before them. This proportion does not seem to jibe with the authors' report, which indicates that only 19 percent and 20 percent of students' problems were "resolved" in Years 1 and 2, respectively. Further, we are never told what "resolved" means. Does it mean that performance deficiencies were eliminated in the short and long term? We have no way of knowing this because no student outcome data were presented. The same may be said of fidelity of treatment data. An absence of fidelity data is all the more noteworthy in light of Flugum and Reschly's (1994) study of the quality of prereferral interventions across Iowa, a forerunner of Heartland's current RTI model. Those interventions, according to Flugum and Reschly, did not reduce the number of students requiring special education because: "(a) Few pre-referral interventions [were] being provided to students; and (b) those . . . that [were] being provided [were] poor in quality" (p. 12).

Referral and placement data, on the other hand, were reported by Ikeda and Gustafson (2002b). But, in the absence of student outcomes, this information is difficult to interpret because referral and placement numbers can be influenced by many administrative and political factors that have little to do with student performance. Finally, Ikeda and Gustafson report that in Year 1, 25 percent of students brought before BATs had been seen previously (p. 10); in Year 2, 34 percent (p. 13). These numbers seem to reflect a large proportion of "false negatives," that is, children whose problems were deemed "resolved" by the BATs but who in fact continued to perform poorly in general education. How much time passed before these children were brought once again before the BATs? Are such figures consistent with the assertion that Heartland's multilevel problem-solving model provides "educational assistance in a timely manner" (Grimes, 2002, p. 8)?

Minneapolis Public Schools

The Minneapolis Public Schools' Problem-Solving Model (PSM) was developed in 1993 (Minneapolis Public Schools, 2001, pp. 1, 3) as a nonbiased method of identification that strengthens teachers' focus on instruction and encourages them to accommodate a greater diversity of students. Thanks to a waiver from the Minnesota State Department of Education, PSM

is currently conducted in all K–8 sites in the district (Minneapolis Public Schools, p. 3). Minneapolis's PSM strongly reflects Heartland's influence in at least four ways. First, it explicitly replaces the psychometric model with RTI as its means of identifying students for special education. To wit: "The student is not declared eligible for special education because of a discrepancy score between IQ and achievement, but because their expected academic performance did not improve as the result of trying a continuum of progressively more intensive regular education interventions" (Marston, 2001, p. 6). Deno, Grimes, Reschly, and Schrag (2001), an external committee constituted by the district to review PSM, offered a more expansive description in their official report: "The PSM requires systematic formative evaluation through repeated assessment of student performance across time. [S]takeholders—teachers, parents, and students—have an opportunity to collaborate in selecting and evaluating the impact of instructional modifications. The evidence produced in this way then becomes the basis for making decisions about a student's special educational needs" (p. 3).

Second, Minneapolis, like Heartland, appears to use a four-level, behavioral problem-solving model (Marston, 2001, Figure 1). Minneapolis teachers identify a struggling student on the basis of achievement data collected systematically and repeatedly on the student and on all children in the school (Level 1); teachers consult informally with colleagues about how to best modify the learning environment to accelerate the student's growth (Level 2); the problem-solving team delineates referral questions, considers causes of the problem, produces intervention alternatives, and directs the implementation and evaluation of the intervention (Level 3); and the problem-solving team considers special education placement (Level 4; see Minneapolis Public Schools, 2001, Table 2). Whereas Heartland's problem-solving team (at Level 3) is staffed by consultants from outside the school, Minneapolis's team comes from within the school building.

Third, Minneapolis's PSM does not make regular use of commercial tests or cut scores to determine (1) whether adequate improvement has been made at the various levels of the PSM and (2) whether a child is eligible for special education. (Although Marston suggests that performance at or below the fifth percentile is indicative of a severe academic discrepancy, Minneapolis does not require practitioners to abide by this number.) Instead, the PSM encourages practitioners to use multiple data sources, similar to Heartland staff's preference of looking for "a convergence of evidence" (Ikeda & Gustafson, 2002a, p. 6), to determine progress and eligibility. Finally, the PSM, like Heartland, takes a noncategorical approach to special education eligibility. The following is taken from "The Problem Solving Model Parent Information Letter" (cf. Minneapolis Pubic Schools, 2001): "If the child needs special education . . . to make progress, we do *not* label the student 'Learning Disabled' or 'Mild Mentally Impaired.' Instead, we will serve him or her as a 'Student Needing Alternative Programming['] (SNAP) for significant academic needs" (emphasis in the original). SNAPs are judged informally to have normal intellectual ability, and their performance in oral expression, listening comprehension, written expression, basic reading skills, reading comprehension, math calculations or math reasoning is discrepant from age-level expectations.

Although the Minneapolis Public Schools (2001) report that, "The problem-solving model has undergone considerable evaluation" (p. 3), we found few published or unpublished evaluations; none in peer-review journals. In one of three studies described as PSM evaluations, schoolwide reading gains for students at PSM sites were significantly higher than for students at non-PSM sites (Minneapolis Public Schools, p. 3). No data were presented, however, on the academic performance of the children (i.e., SNAPs) to whom the PSM was directed. A second study, conducted by Reschly and Starkeweather, explored the quality of interventions for SNAPs and for similar students at non-PSM sites in the Minneapolis Public Schools. According to Marston (2001), the PSM interventions were of superior quality, but "superior" is not explained. Moreover, from Marston's account, it does not seem that Reschly and Starkeweather reported student outcomes.

The third study, described in Marston, Muyskens, Lau, and Canter (in press), was an evaluation of the Minneapolis Public Schools' 2001–2002 kindergarten literacy initiative. The district trained kindergarten teachers to implement specific interventions to boost students' phonological awareness and beginning decoding skills. The teachers were also directed to administer reading assessments in Fall, Winter, and Spring, and to regularly review students' growth in relation to established early-reading benchmarks. In addition, some schools began offering full-day kindergarten classes. In designated underachieving schools, kindergarten teachers were trained to provide intensive small-group instruction to struggling readers. Analyses of student outcomes explored whether the poor readers given intensive interventions in underachieving schools performed better than poor readers, also receiving intensive instruction, in more successful schools. Additional analyses looked at the reading achievement of children in half-day versus full-day kindergartens. Whereas these and other analyses are no doubt important to the district's ambitious effort to strengthen the reading performance of its young children, none seems to represent an evaluation of the PSM.

Across these studies, we could not find evidence bearing on the question: Do students who participate in the PSM make academic progress? Neither did we find information on the nature of the particular interventions produced by the PSM, nor the accuracy with which they were implemented. The previously mentioned team empowered by the district to review the PSM expressed a similar concern. The team's report says the district is in "need . . . of more consistent student data . . . to document the extent of student progress

in interventions implemented in the child's classroom and school" (Deno et al., 2001, p. 10), and that, "Our impression is that the District needs more time to address the substantial staff development that successful implementation of the PSM requires" (Deno et al., 2001, p. 10).

In a different vein, Minneapolis sought to determine whether switching from traditional methods of LD identification to the PSM changed the number and type of students identified for special education. Marston (2001) discusses a 10-year study of child-count rates and concludes that the PSM "did not significantly add students to the special education population" (p. 7). He also says the type of student identified by the PSM and by traditional methods of LD identification was similar. However, evidence from a couple of investigations seems to suggest otherwise. Marston (2001) and Marston et al. (in press) describe one study conducted by a colleague, Dave Heistad, involving 87 traditionally identified LD students and 34 SNAPs. The two groups' reading and math performances on the Northwest Achievement Levels Test were followed across four grades. Whereas Marston et al. report that the groups' scores were "very similar," their Figures 3 and 4 show that the SNAPs' scores were in fact below those of traditionally identified students across the grades. In another study, the school district's case managers were asked to rate large numbers of students with LD, students with mild mental retardation, and SNAPs on personal goals developed by the state. Marston et al.'s Figures 5, 6, and 7 show that the SNAPs were consistently rated lower than the students identified as LD by traditional methods.

This second set of studies raises a couple of questions. First, with respect to the child-count data, shouldn't we expect the PSM to reduce numbers of referrals to and placements in special education? After all, the four-level PSM is meant to provide comparatively intensive and well-designed interventions to preclude the placement of many students in special education. Second, regarding SNAPs versus traditionally identified students, shouldn't the SNAPs perform more poorly because they have demonstrated unresponsiveness to three increasingly intense levels of service? Marston (2001) says "no," but data from his district seem to say "yes." So, too, do findings from a recent pilot study of multilevel problem solving in Ohio (see McNamara & Hollinger, 2002).

Summary

Heartland and Minneapolis, like Ohio and Pennsylvania, use RTI to provide instructional support in a timely manner, and to identify students in need of special education services. However, Heartland's and Minneapolis's RTI model differ in several respects. First, in Ohio and Pennsylvania (and in many other states with mandated prereferral intervention programs), student responsiveness is explored by a teacher, or teacher helper, who implements a time-limited intervention in the mainstream classroom. When a targeted student fails to make adequate progress, she or he is referred to a multidisciplinary team for formal evaluation and possible placement in special education. By contrast, in Heartland and Minneapolis, multiple levels of treatment are available to student and teacher and, as the unresponsive student moves from one level to the next, more resources, including increased expertise, are brought to bear. In other words, Heartland and Minneapolis use a recursive and increasingly intensive prereferral intervention approach, the primary aim of which appears to be remediation, rather than identification. Second, students in Heartland and Minneapolis for whom the multilevel approach is ineffective typically move directly into special education without formal testing by multidisciplinary teams as in Ohio and Pennsylvania. Third, Heartland and Minneapolis reject the use of categorical special education labels. In Heartland, children who receive such services are simply "eligible for special education"; in Minneapolis, they are referred to as "students needing alternative programming," or SNAPs.

Finally, although there is insufficient evidence of the effectiveness of RTI approaches in Ohio and Pennsylvania, this seems especially true for Heartland's and Minneapolis's version. Few published or unpublished studies are available on the latter two sites. These studies have typically involved small (or undefined) samples of schools, teachers, and students, and offer little information about what interventions are implemented, with what degree of accuracy and effectiveness. Moreover, the evaluations do not report how long unresponsive students remain unresponsive before they receive effective remediation in general education or special education services. The absence of such evidence weakens an important assumption among RTI advocates, namely, that RTI provides feasible, timely, and effective interventions. Only feasible, timely, and effective interventions permit one to claim that RTI is preventive for many, and that it distinguishes struggling students with disabilities from others struggling because of inadequate prior instruction.

STANDARD-PROTOCOL APPROACH

A standard-protocol approach to RTI requires use of the same empirically validated treatment for all children with similar problems in a given domain. It may offer several advantages over the problem-solving model: everyone knows what to implement, and it is easier to train practitioners to conduct one intervention correctly and to assess the accuracy of implementation. Assuming it was previously validated by rigorous research, yet another advantage may be that large numbers of students would participate in a generally effective treatment protocol. This approach is advocated by early reading researchers (cf. Lyon et al., 2001) and professional organizations (e.g., International Dyslexia

Association, 2002). It is illustrated by the work of Vellutino and colleagues (e.g., Vellutino et al., 1996), to which we now turn.

Vellutino et al. (1996) asked first-grade teachers to nominate their poorest readers at the beginning of the school year. Those scoring in the lowest 15th percentile on either the Word Attack or Word Identification subtests of the Woodcock Reading Mastery Test-Revised (WRMT-R) became potential study participants. From this group, Vellutino et al. excluded all students who had severe hearing or vision problems, frequent ear infections, and severe emotional problems. The researchers also removed children who were taking daily medication, spoke English as a second language, and had been given a diagnosis of pervasive neurological disorder. Finally, children whose Verbal or Performance IQ scores were 90 or below were also eliminated as study participants.

At the start of the second semester of first grade, the remaining poor readers were assigned 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 amounted to between 70–80 tutoring sessions. Instructors were certified and experienced teachers who had received 30 hours of training in Vellutino et al.'s tutoring program. Tutoring focused on phonemic awareness, decoding, sight word practice, comprehension strategies, and reading connected text. In Fall of second grade, tutored students who were below the 40th percentile on the Basic Skills Cluster participated in an additional 8–10 weeks of tutoring.

As described, between Winter of first grade and Spring of second grade, students were repeatedly administered the Word Attack and Word Identification subtests of the WRMT-R. Responsiveness to intervention was based on slopes derived from linear regression analysis of the WRMT-R data. Vellutino et al. (1996) described four levels of responsiveness: "very limited growth," "limited growth," "good growth," and "very good growth." Two-thirds of the tutored readers demonstrated "good growth" or "very good growth" after one semester of tutoring in first grade. Indeed, they had basically caught up to their normal-reading classmates. Vellutino et al. 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 two subtests of the WRMT-R despite receiving tutoring in both first and second grade. The researchers described these children as "difficult-to-remediate."

We have presented Vellutino et al.'s (1996) work partly because it is representative of researchers' use of a standard-treatment protocol to remediate struggling readers and to identify "nonresponders" (see also Torgesen et al., 1999). In addition, the work illustrates admirable strengths of (and a need to ask questions about) the standard-protocol approach. Vellutino et al.'s intervention, for example, was based on solid experimental work demonstrating its capacity to accelerate young students' reading development. (Yet, this research and similar work of others typically excludes many lowest-achieving readers from participation. It is unclear whether Vellutino and associates' standard protocol is sufficiently robust to help, say, students whose IQ scores are below 90.) Many scholars would use adjectives like "rich" and "comprehensive" to characterize the content of Vellutino et al.'s treatment. (At the same time, one can ask whether their intervention, or any standard treatment, is appropriate for all struggling readers. That is, what may be sacrificed by depriving poor readers of the individually tailored instruction they are more likely to receive in the problem-solving approach?) Finally, Vellutino and his colleagues describe their intervention with sufficient clarity and detail so that others may replicate it. (But how many schools have the resources to provide all their poor readers with 70–80 sessions of one-to-one tutorials conducted by highly trained personnel?)

Summary

In comparing the standard-protocol approach to the problem-solving model, the standard-protocol approach seems more likely in principle to facilitate greater quality control; the problem-solving model appears more sensitive to individual differences. But whereas Vellutino et al. (1996) and other researchers have demonstrated a cause and effect relationship between their standard protocols and improved academic performance, practitioners using problem solving, by and large, have failed to do so. If there are indeed BATs in Heartland and problem-solving teams in Minneapolis that, in concert with classroom teachers, enhance student achievement, then practitioners must document these successes by conducting well-designed evaluations and producing appropriate student outcome data. We know of no practitioners more capable than those in Heartland and Minneapolis of achieving this. Finally, neither the (prereferral intervention and multilevel versions of the) problem-solving model nor the standard-protocol approach has yet proved feasible for large-scale adoption. (The ISTs of Pennsylvania may represent an exception to this statement.) Practitioners of the problem-solving approach typically have not produced fidelity of implementation information or they have documented low levels of implementation accuracy. To our knowledge, the standard-protocol approach has been used exclusively by researchers. It remains to be seen how it will play if adopted by schools. The Bush Administration rightly argues that scientifically-based practices should be used in our schools. Proponents of RTI as an alternative means of LD identification must still prove that their problem-solving approach or their standard-protocol model are worthy of the descriptor "scientifically based."

IMPLICATIONS FOR THE LD CONSTRUCT

Advocates, practitioners, and others with vested interest in the LD construct are facing a dilemma. On one hand, a consensus grows that the IQ-achievement discrepancy should be abandoned as a marker of the disability. On the other hand, as we have tried to make clear, there is an absence of a validated replacement. How to wriggle free from this rock-and-a-hard-place situation? We see three alternatives. We call them "RTI Today," "RTI Tomorrow," and "RTI Never."

Today, Tomorrow, and Never

Supporters of the "RTI Today" option want the immediate replacement of IQ-achievement discrepancy with RTI. They wish to see this substitution not just in the nation's heartland, but everywhere. They brush aside concerns like ours about the inadequacy of current RTI research because they are sure that, however difficult the transition may be to widespread RTI implementation, and irrespective of how imperfect that implementation may be, things cannot get much worse.

Others, who also advocate for RTI but do not share RTI Today proponents' confidence and optimism, see "RTI Tomorrow" as a more prudent strategy. They call for careful and comprehensive evaluations of various RTI approaches. And they support a deliberate scaling up of those that can be implemented with fidelity; that provide demonstrable help in a timely fashion; that identify special-needs students with a limited number of false positives and false negatives; and that reduce special education enrollment, thereby permitting special educators to work intensively with only those students with the most severe learning problems. RTI Tomorrow backers also acknowledge that related issues require discussion. We mention two. First, Heartland's and Minneapolis's version of RTI represents more than a redefinition of LD: it signifies an end to the LD category and construct. Are parents, professional organizations, and others ready for a noncategorical approach to service delivery? Are they willing to combine LD with mild mental retardation and behavior disorders into an undifferentiated "high-incidence-disabilities" entity as some have proposed?

Second, embedded in this categorical versus noncategorical issue is a question of the usefulness of the intelligence construct. Some RTI supporters claim it has no bearing on learning and, hence, should play an equally unimportant role in selecting students for special education and assigning them to instructional groups (cf. Lyon et al., 2001). Others point to more than 50 years of research documenting an empirical relationship between intelligence and school achievement (e.g., Elliott, 1987; Mehrens & Lehmann, 1987). Is there agreement that the routine grouping of children with IQs between 115 and 70, for example, has no important implications for pedagogy?

Finally, the "RTI Never" choice represents those who take a dim view of RTI. They see it as a more ambitious and complex version of prereferral intervention, which many practitioners nationwide have had great difficulty implementing. They ask: What additional personnel, professional development, and leadership will make RTI successful? They also ask a different question: Are we overlooking simpler, psychometric approaches to LD identification that are more defensible than IQ-achievement discrepancy and more likely to be used with fidelity than RTI? (See D. Fuchs, Fuchs, Mathes, Lipsey, & Roberts, 2001, for a description of one such approach.)

Our Preferences

Among the RTI Today, RTI Tomorrow, and RTI Never alternatives, we lean toward RTI Tomorrow. We are also partial to a two-level version of RTI that involves a standard-treatment protocol. The first of the two levels would be a mainstream classroom in which the teacher has been supported by the district to implement research-validated instruction. At the second level, small groups of three to six nonresponsive students would participate in a demonstrably effective standard-treatment protocol.

Our preference for a two-level approach, rather than a three- or four-level model, is based on an impression that fewer levels serve identification better. A student's nonresponsive performance at each of the two levels may be seen as evidence that she requires the individualized, intensive, expert help that special education should provide. By contrast, consider what it would mean if this same student finally responds to instruction at the highest level in a three-tier version; a treatment conducted by a reading specialist 45 minutes per day, five days per week, in a group of three children. Is this finally responsive child ready to return to the classroom? Or, given the intensity of effort and expertise that was required to promote responsivness, is she more appropriately placed in special education? For those answering, "ready to return to the classroom," another question: Is her teacher likely to continue the kind of instruction she received at Level 3? We think not. The higher the level, and the more specialized and intensive the instruction, the greater the disjunction between it and the classroom. The less the likelihood of transfer and generalization of learning and sustained academic achievement.

Another reason to favor a two-level version: because of a well-known inverse relationship between instructional complexity and fidelity of implementation, the greater the number of levels, the less practical RTI becomes. Fidelity concerns are also a basis of our preference for use of a standard-treatment protocol, rather than problem solving, at our second level. It is easier to prepare practitioners to conduct a standard protocol than it is to ready them to implement problem solving because the standard protocol is just that—standard. Practitioners are required to become expert at what is basically one thing. Problem solving requires a more complex set of skills and a greater range of knowledge

because the process is never standard, which is both its strength and weakness.

Finally, we would like to see some valid cognitive assessment administered to all second-level nonresponsive children. This will facilitate identification of students with LD in historical terms, that is, students who fail to learn despite (1) average or better intelligence and (2) participation in generally effective instruction. We are aware of current objections to the use of cognitive assessment, and we are sympathetic to several. Nevertheless, without systematic use of cognitive assessment, the LD construct is on a slippery slope that leads to a bin called "high-incidence disabilities." It will be the end of the LD category.

As indicated, this version of RTI is our preference. We claim no received wisdom and no corner on the truth. With IDEA reauthorization looming, and support weakening for IQ-achievement discrepancy, we urge stakeholders to think imaginatively about RTI and non-RTI alternatives, including what their respective intended and unintended consequences may be; to consider whether adoption of alternative methods should proceed swiftly or deliberately with a role for research; to discuss important issues, such as noncategorical service delivery, implicit in some RTI proposals; and to be mindful about practices that have worked, those that have failed, and what is best for children with LD.

AUTHOR NOTE

Mock, Morgan, and Young are listed alphabetically as authors because they contributed equally to the conceptualization of the article and to the research of various responsiveness-to-intervention approaches. We thank Stan Deno, Tom Jenkins, Joe Kovaleski, Bob Lichtenstein, Doug Marston, Jean Papandrea, Dan Reschly, and Joe Torgesen for sharing documents on RTI, and we are grateful to Deno, Lichtenstein, Marston, Andrea Canter, Jeff Grimes and colleagues, and Sharon Vaughn for written feedback on an earlier draft. Their generous and helpful guidance should not be construed as support for the positions expressed in the article.

REFERENCES

Al Otaiba, S., & Fuchs, D. (2003). *A longitudinal study of children who do not respond to early literacy intervention*. Manuscript in preparation.

Alpert, J. L., & Yammer, D. M. (1983). Research in school consultation: A content analysis of selected journals. *Professional Psychology, 14,* 604–612.

Bergan, J. R. (1970). *Behavioral consultation*. Columbus, OH: Charles E. Merrill.

Bergan, J. R., & Kratochwill, T. R. (1990). *Behavioral consultation and therapy*. New York: Plenum.

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

Canter, A. (2003). Personal communication, February 4.

Chalfant, J. C., Pysh, M. V., & Moultrie, R. (1979). Teacher assistance teams: A model for within-building problem solving. *Learning Disability Quarterly, 2,* 85–96.

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

Conway, S. J., & Kovaleski, J. F. (1998). A model for statewide special education reform: Pennsylvania's Instructional Support Teams. *International Journal of Educational Reform, 7,* 345–351.

Corpolongo, J., Pochowski, A., Lloyd-Jones, A., & Lenz, G. (2002). *Report to the NASDE board: Ad-hoc committee on learning disabilities*. Unpublished paper.

Deno, S., Grimes, J., Reschly, D., & Schrag, J. (2001). *Minneapolis Public Schools—Problem solving model: Review team report*. Unpublished paper.

Dickman, E. G., Hennessy, N. L., Moats, L. C., Rooney, K. J., & Tomey, H. A. (2002). *The International Dyslexia Association response to OSEP Summit on Learning Disabilities*. Unpublished paper.

Division of Learning Disabilities of the Council for Exceptional Children. *Response to OSEP's consensus statement*. Unpublished paper.

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

Elliott, R. (1987). *Litigating intelligence: IQ tests, special education, and social science in the courtroom*. Dover, MA: Auburn House.

Erchul, W. P., & Martens, B. K. (2002). *School consultation: Conceptual and empirical bases and practices* (2nd ed.). New York: Kluwer Academic/Plenum.

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.

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*(1), 6–23.

Flugum, K. R., & Reschly, D. J. (1994). Prereferral interventions: Quality indices and outcomes. *Journal of School Psychology, 32*(1), 1–14.

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.

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88*(1), 3–17.

Frankenberger, W., & Franzaglio, K. (1991). A review of states' criteria for identifying children with learning disabilities. *Journal of Learning Disabilities, 24,* 495–500.

Friend, M., & Cook, L. (1992). *Interactions: Collaboration skills for school professionals*. New York: Longman.

Fuchs, D., & Fuchs, L. S. (1989). Exploring effective and efficient prereferral interventions: A component analysis of behavioral consultation. *School Psychology Review, 18,* 260–283.

Fuchs, D., & Fuchs, L. S. (1996). Consultation as a technology and the politics of school reform. *Remedial and Special Education, 17*(6), 386–392.

Fuchs, D., Fuchs, L. S., & Bahr, M. W. (1990). Mainstream Assistance Teams: A scientific basis for the art of consultation. *Exceptional Children, 57,* 128–139.

Fuchs, D., Fuchs, L. S., Bahr, M. W., Fernstrom, P., & Stecker, P. M. (1990). Prereferral intervention: A prescriptive approach. *Exceptional Children, 56,* 493–513.

Fuchs, D., Fuchs, L. S., Harris, A. H., & Roberts, P. H. (1996). Bridging the research-to-practice gap with Mainstream Assistance Teams: A cautionary tale. *School Psychology Quarterly, 11*(3), 244–266.

Fuchs, D., Fuchs, L. S., Mathes, P. G., Lipsey, M. W., & Roberts, P. H. (2001). *Is "learning disabilities" just a fancy term for low achievement? A meta-analysis of reading differences between low achievers with and without the label*. Paper written for the Office of Special Education Programs, U.S. Department of Education, and presented at the OSEP's LD Summit conference. Washington, DC.

Fuchs, L. S. (1995). *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., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly, 25,* 33–46.

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

Graden, J. L., Casey, A., & Bonstrom, O. (1985). Implementing a prereferral intervention system: Part II. The data. *Exceptional Children, 51,* 487–496.

Grimes, J. (2002). *Responsiveness to interventions: The next step in special education identification, service and exiting decision making*. Revision of paper written for the Office of Special Education Programs, U.S. Department of Education, and presented at the OSEP's LD Summit conference. Washington, DC.

Hallahan, D. P., & Mercer, C. D. (2001). *Learning disabilities: Historical perspectives*. Paper written for the Office of Special Education Programs, U.S. Department of Education, and presented at the OSEP's LD Summit conference. Washington, DC.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2001). *Inferring program effects for specialized populations: Does special education raise achievement for students with disabilities?* Unpublished manuscript.

Hartman, W. T., & Fay, T. A. (1996). Cost-effectiveness of instructional support teams in Pennsylvania. *Journal of Educational Finance, 21,* 555–580.

Horowitz, S. H., Lichtenstein, B., Roller, C., Toombs, M., & Kukic, S. (2002). *An intervention-oriented, multi-tier approach for identifying and serving students with learning disabilities*. Unpublished paper prepared for the Learning Disabilities Roundtable, "Finding Common Ground."

Ikeda, M., & Gustafson. (2002a). *Heartland AEA 11's problem solving process: Impact on issues related to special education* (Research Report No. 2002-01). Available from authors at Heartland Area Education Agency 11, 6500 Corporate Dr., Johnston, IA, 50131.

Ikeda, M., & Gustafson. (2002b). *Heartland AEA 11's problem solving process: Impact on issues related to special education* (Research Report No. 2002-01). Available from authors at Heartland Area Education Agency 11, 6500 Corporate Dr., Johnston, IA, 50131.

Ikeda, M. J., Tilly, D. W., Stumme, J., Volmer, L, & Allison, R. (1996). Agency-wide implementation of problem-solving consultation: Foundations, current implementation, and future directions. *School Psychology Quarterly, 11*(3), 228–243.

Kaufman, M. J., Kameenui, E. J., Birman, B., & Danielson, L. (1990). Special education and the process of change: Victim or master of educational reform? *Exceptional Children, 57*(2), 109–115.

Kavale, K. A. (1987). Theoretical issues surrounding severe discrepancy. *Learning Disabilities Research, 3*(1), 12–20.

Kovaleski, J. F. (2002). Best practices in operating pre-referral intervention teams in Pennsylvania. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-IV* (pp. 645–655). Bethesda, MD: National Association of School Psychologists.

Kovaleski, J. F., Gickling, E. E., Morrow, H., & Swank, H. (1999). High versus low implementation of instructional support teams. *Remedial and Special Education, 20*(3), 170–183.

Kratochwill, T. R., Elliott, S. N., & Rotto, P. C. (1995). Best practices in school-based behavioral consultation. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-III* (pp. 519–535). Washington, DC: National Association of School Psychologists.

Kratochwill, T. R., & Van Someren, K. R. (1985). Barriers to treatment success in behavioral consultation: Current limitations and future directions. *Journal of School Psychology, 23,* 225–239.

Lyon, G. R., Fletcher, J. M., Shaywitz, S. E., Shaywitz, B. A., Torgesen, J. K., Wood, F. B. et al. (2001). *Rethinking learning disabilities*. Hudson Institute.

MacMillan, D. L., Gresham, F. M., & Bocian, K. M. (1998). Discrepancy between definitions of learning disabilities and school practices: An empirical investigation. *Journal of Learning Disabilities, 31*(4), 314–326.

Marston, D. (2001). *A functional and intervention-based assessment approach to establishing discrepancy for students with learning disabilities*. Paper written for the Office of Special Education Programs, U.S. Department of Education, and presented at the OSEP's LD Summit conference. Washington, DC.

Marston, D., Muyskens, P., Lau, M., & Canter, A. (in press). Problem solving model for decision-making with high-incidence disabilities. *Learning Disabilities Research & Practice*.

McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2003). *Identification of treatment nonresponders to generally effective reading instruction: An experimental field trial*. Manuscript in preparation.

McNamara, K., & Hollinger, C. (2002). Intervention-based assessment: Evaluation rates and eligibility findings. *Exceptional Children, 69*(2), 181–193.

Medway, F. J., & Updyke, J. F. (1985). Meta-analysis of consultation outcome studies. *American Journal of Community Psychology, 13,* 489–505.

Mehrens, W. A., & Lehmann, I. J. (1987). *Using standardized tests in education* (4th ed.). New York: Longman.

Mercer, C. D., King-Sears, P., & Mercer, A. R. (1990). Learning disabilities definitions and criteria used by state education departments. *Learning Disability Quarterly, 13,* 141–152.

Meyers, J., Pitt, N. W., Gaughan, E. J., & Friedman, M. P. (1978). A research model for consultation with teachers. *Journal of School Psychology, 16,* 137–145.

Minneapolis Public Schools. (2001). *Problem solving model: Introduction for all staff* . Minneapolis, MN: Author.

Morris, R. D., Shaywitz, S. E., Shankweiler, D. P., 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 School Psychologists. (2002). Draft: *NASP recommendations on LD eligibility and identification for IDEA reauthorization*. Unpublished paper.

National Center for Learning Disabilities. (2002). NCLD calls for national mandate on early intervention. *Our World, 7*.

Office of Special Education Programs. (2002). *Specific learning disabilities: Finding common ground*. Washington, DC: Author.

Papandrea, J., Walkley, B., DeLorme, L., & McNamara, D. (2001). *The 7 Share Initiative: Year 4 program evaluation*. Unpublished paper.

Ponti, C. R., Zins, J. E., & Graden, J. L. (1988). Implementing a consultation-based service delivery system to decrease referrals for special education: A case study in organizational considerations. *School Psychology Review, 17,* 89–100.

President's Commission on Excellence in Special Education. (2002). *A new era: Revitalizing special education for children and their families*. Washington, DC: Author.

Pugach, M. C., & Johnson, L. J. (1989). Prereferral interventions: Progress, problems, and challenges. *Exceptional Children, 56,* 217–226.

Pugach, M. C., & Johnson, L. J. (1995). *Collaborative practitioners, collaborative schools*. Denver, CO: Love.

Reynolds, M. C., Wang, M. C., & Walberg, H. J. (1987). The necessary restructuring of special and regular education. *Exceptional Children, 53,* 391–398.

Rutter, M., & Yule, W. (1975). The concept of specific reading retardation. *Journal of Child Psychology and Psychiatry, 16,* 181–197.

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.

Shepard, L., & Smith, M. L. (1983). An evaluation of the identification of learning disabled students in Colorado. *Learning Disability Quarterly, 6,* 115–127.

Sheridan, S. M., Welch, M., & Orme, S. F. (1996). Is consultation effective? A review of outcome research. *Remedial and Special Education, 17*(6), 341–354.

Siegel, L. S. (1989). I.Q. is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities, 22,* 469–478.

Siegel, L. S. (1999). Issues in the definition and diagnosis of learning disabilities: A perspective on Guckenberger v. Boston University. *Journal of Learning Disabilities, 32,* 304–319.

Skrtic, T. M. (1991). *Behind special education: A critical analysis of professional culture and school organization*. Denver, CO: Love.

Sleeter, C. E. (1998). Yes, learning disabilities is political; what isn't? *Learning Disability Quarterly, 21,* 289–296.

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., Case, L. P., & Molloy, D. E. (in press). Responsiveness to general education instruction as the first gate to learning disabilities identification. *Learning Disabilities Research & Practice*.

Speece, D. L., & Shekitka, L. (2002). How should reading disabilities be operationalized? A survey of experts. *Learning Disabilities Research & Practice, 17*(2), 118–123.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360–406.

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*(3), 443–461.

Tharp, R. G., & Wetzel, R. J. (1969). *Behavior modification in the natural environment*. New York: Academic Press.

Torgesen, J. K., Morgan, S. T., & Davis, C. (1992). Effects of two types of phonological awareness training on word learning in kindergarten children. *Journal of Educational Psychology, 84,* 364–370.

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

Vaughn, S., & Linan-Thompson, S. (2003). *What's special and effective in reading instruction for students with LD*? Paper presented at the Eleventh Annual Pacific Coast Research Conference. La Jolla, CA.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., 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.

Viadero, D. (1991). States turn to special education programs for budget cuts. *Education Week,* June 12, p. 16.

West, J. F., & Idol, L. (1987). School consultation (part 1): An interdisciplinary perspective on theory, models, and research. *Journal of Learning Disabilities, 20,* 388–408.

Witt, J. C., Gresham, F. M., & Noell, G. H. (1996). What's behavioral about behavioral consultation? *Journal of Educational and Psychological Consultation, 7*(4), 327–344.

Ysseldyke, J. E., Algozzine, B., & Epps, S. (1983). A logical and empirical analysis of current practices in classifying students as learning disabled. *Exceptional Children, 50,* 160–166.

About the Authors

Douglas Fuchs is a Professor of Special Education at Peabody College of Vanderbilt University and Co-Director of the National Center for Research on Learning Disabilities, a project jointly run by Vanderbilt and the University of Kansas and funded by the Office of Special Education Programs in the U.S. Department of Education.

Devery Mock is completing her Ph.D. from the University of Virginia. In the Fall of 2003, she will be an Assistant Professor of Special Education at the University of Iowa. Her research interests include developing effective practices for identifying students with learning disabilities.

Paul Morgan is a doctoral student at Peabody College of Vanderbilt University. In Winter, 2004, he will become an Assistant Professor of Special Education at Pennsylvania State University. His research interests include the early identification and remediation of reading disabilities.

Caresa Young is completing her Ph.D. at Peabody College of Vanderbilt University. She is an Assistant Professor of Dyslexia Studies at Middle Tennessee State University, where she teaches classes and conducts research on early reading problems.