Identifying Students at Risk, Monitoring Performance, and Determining Eligibility Within Response to Intervention: Research on Educational Need and Benefit From Academic Intervention

Mark R. Shinn, PhD
National Louis University, Skokie, Illinois

Abstract. The new Individuals With Disabilities Education Improvement Act (2004) offers local education agencies the choice of using a student’s response to intervention (RTI) as a major component to determine eligibility for special education under the category of specific learning disabilities (SLD). Using a RTI model, it is not expected that different students will be identified as SLD than those identified historically. For more than 25 years, accumulated evidence has strongly suggested that most students labeled SLD have been those students with severe educational needs (i.e., have performance discrepancies compared to students in their own communities), regardless of the stated eligibility criterion (e.g., ability–achievement discrepancies). What is unique about RTI is that educational need is a necessary but not sufficient requirement for SLD identification. Students also must not be responding to high-quality general education instruction (i.e., receiving educational benefit). This article first explores the scientific evidence for operationalizing the educational need component using curriculum-based measurement. Second, the use of curriculum-based measurement to assess RTI is presented not only within a special education eligibility process, but also for universal screening and progress monitoring for all students within a three-tier problem-solving model. The article concludes with important research questions within RTI.

Recent changes in federal special education law resulted in a dramatic reconceptualization of the process that educators could use to identify a student as eligible for special education under the category of specific learning disabilities (SLD). As discussed in more detail in previous articles in this special issue, local education agencies have been given the option of using a process to determine whether a student responded to high-quality, scientifically based intervention. What is often surprising to many practitioners is that the response to intervention (RTI) process was offered as a potential remedy to an eligibility process based primarily on computing the discrepancy between ability and achievement, and to a

Financial disclosure: Mark R. Shinn serves as a consultant to The Psychological Corporation in his position as chief scientist for AIMSweb. AIMSweb produces curriculum-based measurement testing materials and organizes and reports results in a three-tier model, including response to intervention.

Correspondence regarding this article should be addressed to Mark R. Shinn, National Louis University, Skokie, IL 60077; E-mail: markshinn@mac.com

Copyright 2007 by the National Association of School Psychologists, ISSN 0279-6015
much less extent, on identifying a processing deficit.

In this article I will briefly review more than 20 years of research revealing that what has been considered a “traditional approach” (i.e., assessing the ability–achievement discrepancy) is an insufficient explanation for the millions of students who have been labeled as SLD. Using a deductive approach, significant numbers of students identified as SLD fail to meet the stated eligibility standards such that these approaches must be rejected as an explanation of school practices. I also will examine the research on school-based identification of students as SLD from an inductive approach that resulted in highly consistent SLD status markers. I then will detail how these last research findings can be translated into time-efficient and accurate SLD identification practices in an RTI, dual-discrepancy process (Fuchs, 2003; Fuchs, Fuchs, & Speece, 2002), where educational need and educational benefit become the focus of entitlement decision making. Finally, I will present the RTI process within a larger problem-solving framework of a three-tier model wherein all students’ educational needs and educational benefit are assessed in an ongoing manner.

Deductive Research Fails to Confirm Consistent Use of SLD Eligibility Criteria

Because the ability–achievement discrepancy has been the cornerstone of SLD identification since passage of the Education for All Handicapped Children Act in 1975, there is a tendency for practitioners to believe this approach is both well accepted and scientifically based. However, as noted by Algozzine and Ysseldyke (1987) 20 years ago:

The term learning disability was scarcely off the breath of early pioneers when a profession began questioning its own integrity. Today, the proliferation of students classified as learning disabled (LD) has caused social, political, economic, and educational concerns that, in turn, have produced serious questioning of practices (p. 307).

As a point of fact, federal interest in SLD identification and other special education issues began shortly after passage of Education for All Handicapped Children Act regulations in 1977 when five learning disabilities research institutes were funded to investigate potential problems and provide viable solutions. One of the funded institutions, the University of Minnesota Institute for Research on Learning Disabilities, was charged with the study of the assessment and decision-making processes for students with SLD, including how schools operationalized the federal definition and the congruence between SLD eligibility criteria. James Ysseldyke was the principal investigator in this line of research.

In the early 1980s a number of studies emerged from Ysseldyke and the Institute for Research on Learning Disabilities that used a deductive approach to determine if the data collected and used by schools to identify students as SLD matched the severe discrepancy and/or processing criteria. I classify this type of research as “do as I say” research. For example, Ysseldyke, Algozzine, Shin, and McGue (1982) found few reliable differences among students identified as SLD and low-achievers (LA) in their ability–achievement discrepancies or processing performance among a variety of ability, achievement, and other measures. This study was described by Gresham, MacMillan, and Bocian (1996) as “perhaps the most influential and cited study in this debate” (p. 571). The scores of both groups in the Ysseldyke, Algozzine, Shin et al. (1982) study were similar and overlapped more than 90%.

Other studies by this research group yielded similar conclusions (e.g., Algozzine, Ysseldyke, & Shin, 1982; Ysseldyke, Algozzine, & Epps, 1983; Ysseldyke, Algozzine, Richey, & Graden, 1982) and were summarized by Ysseldyke, Tuchlow, Graden, Wesson, Algozzine, and Deno (1983):

1. Schools collected lots of data as part of the SLD eligibility process that were related to the eligibility criteria (e.g., ability and achievement scores).
2. Schools failed to use eligibility criteria in a consistent and predictable way. As often as not, schools ignored the data.
3. Based on the testing data collected as part of the eligibility process, it was difficult to detect meaningful differences between students labeled SLD and LA.

4. Schools often identify students as SLD when they do not meet the stated eligibility criteria.

Over a period of more than 25 years, similar outcomes have been reported by other researchers (e.g., Cone, Wilson, Bradley, & Reese, 1985; Gresham, 2002; Gresham et al., 1996; MacMillan & Siperstein, 2002; Shepard & Smith, 1983; Wilson, 1985). More recently, Peterson and Shinn (2002) studied the match between the stated ability-achievement discrepancies of 48 fourth-grade students labeled SLD in high- and low-achieving communities in Minnesota. Using the state prescribed criterion of −1.76 standard deviation discrepancy, corrected for regression, 40% of students did not meet this discrepancy standard.

**Inductive Research Examines What Schools Do**

One of the tempting conclusions from the “do as I say” research is that school SLD identification is a haphazard, almost capricious process because there has been no evidence of systematic application of explicit SLD eligibility criteria. Despite SLD criteria, students labeled SLD frequently do not meet them. Running parallel to this type of research has been *inductive research* that has tried to identify discernable patterns in assessment results. I classify this type of research as “do as I do” research.

The question has been, “Is there an identifiable and systematic pattern of performance in SLD students that adequately explains why schools label them this way?” One common observation within the “do as I say” research studies was that, not surprisingly, school-identified SLD students performed significantly less well on achievement measures than typically achieving students, especially in reading. Notably, however, those achievement differences also were apparent when comparing SLD students with their low-achieving counterparts. For example, in a reanalysis of the Ysseldyke, Algozzine, Shinn et al. (1982) results, Kavale, Fuchs, and Scruggs (1994) reported an effect size difference of −0.34 overall, considered by Cohen (1988) to be a medium effect size. Differences in Pennsylvania’s Initiative on Assistive Technology achievement scores accounted for differences in approximately 80% of the cases of SLD and low-achieving students.

More obvious differences among SLD and LA in reading achievement were noted in the Ysseldyke, Algozzine, Shinn et al. (1982) sample in a follow-up study that formed my dissertation (Shinn, 1981) and was published several years later (Shinn, Ysseldyke, Deno, & Tindal, 1986). When 71 of the original 99 students were compared on curriculum-based measurement (CBM) tests of general reading ability, the differences in the number of words read correctly in 1 min by SLD and LA students were compelling. When an effect size was computed, SLD students read −1.3 standard deviations lower.

**Severe Low Achievement as an Explanation of “What Schools Do” to Identify SLD**

After analyzing the research studies including students that school personnel actually labeled as SLD, severe achievement alone became a compelling explanation for identification practices that was tested systematically. Testing a hypothesis put forth by Gerber and Semmel (1984) that the special education process began by teachers referring the lowest performing students in their classrooms for potential eligibility, Shinn, Tindal, and Spira (1987) studied the reading performance of 570 students in Grades 2–6 referred for potential learning disabilities. Students were tested using reading CBM (R-CBM; Deno, 1989; Shinn, 1989, 1998) procedures, and their scores were compared to the oral reading scores of typical students in the community. Across grades, a consistent pattern was noted. The typical student referred for potential learning disabilities performed below the 5th percentile compared to general education
peers. When converted to an effect size statistic, the difference in R-CBM scores ranged from −1.0 to −1.6 with a median effect size of −1.4. This effect size translates into a large effect (Cohen, 1988).

The highly consistent teacher referral practices led to two studies in the same school district regarding who was actually placed in special education for the category of SLD. The stated policy in the district was that SLD was defined as a severe ability-achievement discrepancy. However, both studies tested the hypothesis that what schools actually did was sort students by the severity of their achievement problems. Students with achievement problems that were not as severe were predicted to be students who were placed in remedial programs. Their scores would be lower compared to typical general education students, but not as low as students placed in special education as LD. It was predicted that students actually placed in special education, regardless of the policy of severe ability-achievement discrepancies, would consistently be the lowest-performing students.

At the level of a specific school, Shinn and Marston (1985) compared the R-CBM scores of 58 SLD students in Grades 4–6 to randomly sampled students in a remedial program (i.e., Title I) and general education. The median effect size difference between typically achieving students and students labeled SLD of −2.3 was consistent (range −2.3 to −2.4) and large. The median effect size difference between LA students in Title I and students labeled SLD of −1.1 also was consistent (range −1.0 to −1.3) and large.

At the level of a school district, Shinn, Tindal, Spira, and Marston, (1987) compared the R-CBM scores of 638 SLD students in Grades 1–6 to 451 randomly sampled students in a remedial program (i.e., Title I) and 2337 general education students. The typical student labeled SLD performed at the 3rd percentile across grades. Performance of students in the Title I remedial program was also highly consistent with the prediction. Although these students were low, a median percentile of approximately the 16th percentile, their scores were markedly higher than students labeled LD. In this study, the median effect size difference between typically achieving students and students labeled SLD of −1.85 was consistent (range −1.1 at Grade 1 to −2.4 at Grade 4) and large. The median effect size difference between LA students in Title I and students labeled SLD of −0.95 also was consistent (range −0.2 at Grade 1 to −1.3 at Grade 6) and, with the exception of Grade 1, was large.

Studies by other authors confirmed these findings using other methods, populations, and measures. For example, Gresham et al. (1996) reported effect sizes of −0.64 between SLD and LA students in a sample of 152 second-through fourth-graders with what they described as a differentiation rate of 74%. Finally, a recent meta-analysis of 79 studies comparing SLD and low-achieving students (Fuchs, Fuchs, Mathes, Lipsey, & Eaton, 2000; Fuchs, Fuchs, Mathes, Lipsey, & Roberts, 2001) lends strong support to the defining feature of SLD as being severe low achievement. The largest and most consistent difference between school-identified SLD students and low-achieving students was extreme low achievement, with an effect size of 0.61. The authors reported that “73% of the LA population fell above the average reading score of the LD population” (p. 4).

Building Effective Identification Practices for RTI

If schools want to serve students with similar characteristics as they have been serving for more than 25 years, then it would make sense to identify students with severe achievement needs as potentially eligible for special education as SLD. Students labeled SLD must be significantly different from other students on the target academic skill domain (e.g., reading). This approach is consistent with a portion of the recommendation the National Research Council (2002) in their report on disproportional placement in special education. They recommended, in part, that “eligibility ensues when a student exhibits large differences from typical levels of performance in one or more domain(s)” (p. 6). A hypothet-
ical illustration of this achievement discrepancy is shown in Figure 1.

A severe discrepancy criterion (i.e., cutting score) is established relative to typical levels of achievement. Students below the criterion may be considered to have severe educational need (Shinn, 2005b, in press), the first component of a dual-discrepancy approach to response to intervention (Fuchs et al., 2002; Pericola Case, Speece, & Eddy Molloy, 2003; Speece, Pericola Case, & Eddy Molloy, 2003). To translate previous SLD identification research into SLD identification practice within RTI, two questions must be answered:

1. How large a discrepancy from “typical levels of achievement” is required?
2. What constitutes “typical achievement,” and to whom should a student be compared to determine the magnitude of the discrepancy?

**How Large Should the Discrepancy Be?**

Given the observed reading achievement discrepancies in published studies between SLD students and typically achieving students, it seems logical that achievement at or below the 5th percentile could be defended as one that would identify students consistent with school-based SLD identification practices. For example, in Shinn, Tindal, Spira et al. (1987), the overarching characteristic of students placed in SLD programs was reading scores at or below the 5th percentile, with a median of the 3rd percentile across grades. This criterion also would align well with a projected cut score extrapolated from the Fuchs et al. (2001) meta-analysis comparing SLD with LA students. If a local education agency wanted to ensure that students with potentially severe reading problems were not missed (i.e., false negatives) then the criterion for a severe discrepancy could be adjusted upwards. A cut score of the 10th percentile would be consistent with the reading discrepancies observed when teachers refer students for potential learning disabilities services (Shinn, Tindal, & Spira, 1987).

What Defines Typical Achievement?

As potentially controversial as the achievement discrepancy alone may be, it pales in comparison to the issues raised about the normative comparison group. Unfortunately, neither the National Research Council recommendation (2002) nor the extant research resolves the question of what defines typical achievement sufficiently. The intuitive answer to the question would be that students identified as SLD be significantly different from typically developing peers from a national norm. This definition of a severe achievement discrepancy is reflected in Figure 1 and has been called an absolute achievement discrepancy (AAD) approach. See Peterson and Shinn (2002) and Shinn (2005b) for more detail.

The AAD approach is consistent with the published research on school-based SLD identification. As I have noted, research has established large achievement effect size differences among SLD students, LA students, and typically achieving students. A group of SLD students has been compared to LA and/or typically achieving students on a variety of measures, including tests that have national norms. However, the groups’ scores were compared to each other, not compared to differences from a national norm.

My difficulty with the AAD model is that, although it may do a better job of explaining school-based identification practices than the ability–achievement discrepancy model, it, too, seems logically insufficient. Hypothetically, if a national norm were used to define typical achievement and thus the performance discrepancy, then large numbers of students would be identified as SLD in low-performing communities; few students would be identified as SLD in high-performing communities. I have noted concerns about this issue in previous works. "If the achievement-only discrepancy model were a satisfactory explanation for school decision-making practices, one would predict that there would be a relation between the severity of achievement problems in communities and states and the number of students identified as eligible for special education" (Shinn, 2005b, p. 229). High-performing school districts would have the lowest numbers of students identified as LD. Low-performing school districts would have the highest number of students identified as LD.

No such evidence has been presented in the research literature. As school psychology practitioners are well aware, there is an equal or often even higher proportion of students labeled SLD in high-performing communities as in low-performing communities. In addition, often when students move to high-performing communities from low-performing communities, they are referred for special education as SLD within a short period of time.

A Relative Achievement Discrepancy May Work Best

Only one published study has examined the role of what norm group defines typical achievement in explaining school-based SLD identification. This study, by Peterson and Shinn (2002), compared the classification accuracy of national achievement and local achievement norms with the mandated state ability–achievement discrepancy approach. Students identified as SLD by high- and low-achieving school districts were compared. As described earlier in this article, the ability–achievement discrepancy approach did not explain practice well in either school district. A national norm approach worked reasonably well for students in the low-performing school district. Using the Woodcock-Johnson Broad Reading Cluster (Woodcock & Johnson, 1989), approximately 80% of students with SLD from the low-achieving context were significantly below a national norm. However, a nationally normed definition of typical achievement worked very poorly in the high-achieving school district. Only 22% of students with SLD from the high-achieving school district had severe achievement discrepancies from the national norm. Students placed in special education SLD programs in the high-achieving school district read more like typically achieving students in the low-achieving district. The SLD students in the
high-performing district performed more than a standard deviation better than SLD students in the low-achieving district whether assessed by the Woodcock-Johnson Broad Reading Cluster or R-CBM.

In the Peterson and Shinn (2002) study, when the criterion for a severe achievement discrepancy was severe discrepancy deviations from typical levels of achievement from students from the same community, nearly all students identified as SLD (85–95%) from both communities met this standard. This type of discrepancy has been called the relative achievement discrepancy (RAD; Peterson & Shinn, 2002) and is shown in Figure 2.

A RAD approach, where a performance discrepancy is computed compared to local peers, best explains the results in Peterson and Shinn (2002). This approach also arguably fits the data in previous research when SLD and LA and typically achieving students are compared within the same communities. Notably, it provides a defensible explanation for what schools do and what practitioners observe when students' skill levels do not match those in their community. A full discussion of the role of the norm in defining typical achievement expectations is beyond this article, but the reader is referred to Deno (1989, 1995, 2002) and Shinn (2005b) for more information.

**Severe Educational Need Is Not Enough for RTI**

Regardless of which norms are used to operationalize a performance discrepancy (i.e., severe educational need), the discrepancy typically is a necessary but not sufficient condition for eligibility determination in an RTI approach. Most commonly, a dual-discrepancy approach to RTI (Fuchs et al., 2002; Pericola-Case et al., 2003) is used, whereby (a) severe educational need is combined with (b) the lack of educational benefit (i.e., lack of response) from high-quality intervention. This dual discrepancy was expressed cogently by


607
the National Research Council (2002), who proposed that "eligibility ensues when a student exhibits large differences from typical levels of performance in one or more domain(s) and with evidence of insufficient response to high-quality interventions in the relevant domain(s) of functioning in school settings" (p. 6).

In a dual-discrepancy approach, students may have a severe performance discrepancy (e.g., below the 5th or 10th percentile compared to a local or national peer) but may not be eligible for special education if they are progressing satisfactorily when provided with high-quality general education instruction. Of course, the second component of the dual-discrepancy approach requires progress monitoring. I would argue that the quality of the progress monitoring would be no less than the quality of the intervention; in other words, it too should be scientifically based.

**CBM Fits the Needs of RTI for Monitoring Response**

Because the progress-monitoring component of RTI should employ tools that are scientifically based, it seems logical that CBM (Deno, 1985, 1986, 2003) would be the primary tool in an RTI process when there are concerns about a student's basic skills. CBM was developed more than 25 years ago through federally funded research to provide special education teachers with a way to write objective Individualized Education Program (IEP) goals and continuously monitor progress for those students who ultimately received special education. Given its documented utility in monitoring progress for students with disabilities (Deno, 2003; Fuchs & Fuchs, 1986b, 1999; Fuchs & Vaughn, 2005), it seems appropriate that it be used to monitor progress as part of the RTI disability determination process.

CBM (Deno, 1985, 1986, 1989; Fuchs & Deno, 1991; Shinn, 1989, 1998) is a set of standardized and validated short-duration tests 1–4 min long in reading (e.g., oral reading, maze), mathematics computation, mathematics applications, spelling, written expression, early literacy (e.g., Dynamic Indicators of Basic Early Literacy Skills, DIBELS; Good, Wallin, Simmons, Kame'enui & Kaminski, 2002), and early numeracy (Clarke & Shinn, 2004).

The vast majority of the CBM research has been in the area of reading, whereby it has been demonstrated repeatedly that having students read aloud from text for 1 min and counting the number of words read correctly serves as a highly accurate and efficient measure of general reading ability (Deno, Mirkin, & Chiang, 1982; Fuchs, Fuchs, & Deno, 1982; Fuchs, Fuchs, & Maxwell, 1988; Shinn, Good, Knutson, Tilly, & Collins, 1992). R-CBM and members of the CBM “family” (e.g., DIBELS) have been reviewed and approved for use in progress monitoring by Reading First (Kame'enui, 2002) and the National Center on Student Progress Monitoring, a U.S. Department of Education, Office of Special Education Programs, project (see http://www.studentprogress.org). The Center’s mission is to provide technical assistance to states and districts and disseminate information about progress-monitoring practices proven to work in different academic content areas, Grades K–5.

Using R-CBM, a student’s response to intervention can be evaluated reliably in as little as 4–6 weeks (Fuchs & Vaughn, 2005) should a sufficient number of data points be collected (Fuchs & Vaughn, 2005; Shinn, Good, & Stein, 1989). An illustration of an RTI progress-monitoring graph for a second-grade student, Sam, is shown in Figure 3.

The solid line represents the expected rate of progress that was specified as representing adequate progress for Sam, a student referred for concerns about his discrepancy from peers and his lack of reading progress. The dashed line represents Sam’s actual rate of progress when he was provided a modified reading program after the initial referral meeting. His actual rate of progress exceeded the expected rate of progress and it was decided that the intervention was effective. He responded to the intervention.

A second illustration of an RTI progress monitoring graph for a second-grade student,
Samantha, is shown in Figure 4. Her actual rate of progress was below the expected rate of progress and it was decided that the intervention was ineffective. She did not respond to the intervention and the team may decide to modify the intervention or pursue the need for an IEP.

**CBM: Monitoring Progress in RTI and More**

The documented usefulness of CBM to monitor progress accurately and efficiently makes it an appropriate choice for measuring a specific student’s response to intervention as part of an eligibility process for learning disabilities. However, CBM provides schools with the ability to do more than monitor progress for students as part of RTI. Should a student be eligible for special education, it provides a “best-practice” approach for writing IEP goals and monitoring the effects of the special education intervention (Fuchs & Shinn, 1989; Shinn & Shinn, 2000). It also has been demonstrated to result in increased achievement of students with disabilities (Fuchs & Fuchs, 1986a; Fuchs & Fuchs, 2004). The use of the same progress-monitoring tool before and after special education placement provides a continuous database, increasing the likelihood of understanding the data for all educators and parents, and reducing the training needs for different assessment systems.

CBM brings two more value added benefits to practice when it is included in an RTI process, the ability to (a) measure the performance discrepancy using a local normative approach, and (b) feasibly implement a preventive three-tier basic skills progress monitoring.

**Using CBM to Measure the Performance Discrepancy**

From a RTI perspective, it seems like any quality achievement test could be used to measure the achievement discrepancy from typical peers. However, if the RAD best explains what schools do in identifying students as LD, then it seems as if whatever achievement test is employed should be normed locally.

Shortly after initial validation studies began showing the usefulness of CBM as a progress-monitoring tool in the early 1980s, schools began to use it as a norm-referenced test to determine the performance discrepancy, with referred students as part of the eligibility determination process (Marston, Deno, & Tindal, 1984; Marston & Magnusson, 1985; Tindal, Wesson, Deno, Germann, & Mirkin, 1985). Local norms were used in this process and resulted in more accurate predictions of referral to placement outcomes (Marston et al., 1984). For more than 25 years, CBM has been an integral tool in eligibility determination in a problem-solving model, whereby it is the severe discrepancy from peers from the same

---

**Figure 3.** Response to intervention after 5 weeks for Sam, a Grade 2 student, tested in reading standard progress monitoring passages. (Copyright © 2006 Edformation, Inc. Reprinted with permission.)
community that defines and validates the problem. Illustrations of three potential performance discrepancies are shown in Figure 5.

To measure a performance discrepancy in reading, students typically read three passages and the median score is graphed and compared to peers using a “box-and-whisker” chart. This process takes approximately 5 min. The box represents the range of typically achieving students (i.e., scores between the 25th and 75th percentile). The lower whisker represents scores from the 25th to the 10th percentile. Conversely, scores in the upper whisker represent scores from the 75th to the 90th percentile. Scores below the lower whisker are below the 10th percentile and could meet a definition of a severe performance discrepancy as part of an RTI process. In the three examples juxtaposed in Figure 5, the first student would not have a significant performance discrepancy that would potentially require an RTI process. The student’s reading scores are well above average. The second student may have a performance discrepancy that requires an intervention. This student may be an appropriate candidate for a Tier 2 intervention. The third student’s R-CBM score is well below the 10th percentile and may be a good candidate for RTI.

A number of advantages appear obvious when CBM is used to measure the performance discrepancy component of the RTI dual-discrepancy process. First, the assessment process is accurate and time efficient. Second, the process has a lengthy history of use in school districts that employ a problem-solving model, which generally results in increased service delivery satisfaction for practitioners and parents (Ikeda et al., 2002; Tilly & Grimes, 1998). Third, it comprises a continuous assessment process that uses data collected in similar ways for different assessment purposes.

Using CBM in a Prevention Model to Assess RTI for All Students

A RTI process can be used only as a replacement process for ability–achievement discrepancies in the special education eligibility procedure. That is, students can be referred, tested, albeit on some different dimensions, and placed in special education. However, nearly all major RTI school-based initiatives (Batsche et al., 2005; Fuchs & Vaughn, 2005; Pericola Case et al., 2003; Vaughn & Fuchs, 2003) embed the RTI decision in a variation of a problem-solving model (Deno, 2002; Shinn, in press) that has the following components:

1. Use of a three-tier heuristic to design multiple levels of interventions with attention to evidence-based practices, and the creation of highly effective remedial interventions (e.g., Tier 2) to support the needs of at-risk students.

![Figure 4. Response to intervention after 5 weeks for Samantha, a Grade 2 student, tested in reading standard progress monitoring passages. (Copyright © 2006 Edformation, Inc. Reprinted with permission.)](image-url)
2. A universal screening process to promote early identification and intervention and to reduce the need for individual student referrals.

3. Use of scientifically based progress monitoring to assess RTI for all students.

I prefer to distinguish this conception of RTI as an “rti” process whereby all students’ progress is monitored to ensure they are benefiting from the instruction they are receiving, and if they are not, then changes in the intervention are made. I distinguish this “rti” process from the “RTI” eligibility process, which has the purpose of making a special education entitlement decision because the student has failed to respond within an “rti” process.

Integral to this rti conception is the use of CBM in a three-tier model in a way that combines universal screening and progress monitoring. A full discussion of the use of CBM in a three-tier model is beyond the scope of this article; the reader is referred to Shinn (in press), or Shinn, Shinn, Hamilton, and Clarke, (2002) for more detail. In brief, all students in general education are tested three to four times per year (i.e., benchmark assessment) and an individual student’s scores and rates of improvement are compared to those of peers. An illustration in reading is shown in Figure 6.

The second-grade student in Figure 6 was below average on fall benchmark assessment, and although the student improved somewhat by the time of the winter benchmark testing, the achievement gap was growing. Based on this gap, an increasing performance discrepancy, the student received a remedial Tier 2 intervention in addition to core general education instruction. By the end of the year, at the spring benchmark, the student had improved significantly, and the performance gap was eliminated. The student “responded to the intervention.”

By having all students’ progress monitored, changes can be made early and the effects can be judged. In Figure 7, a student with a severe performance discrepancy was identified, and general education instruction was supplemented with an intensive Tier 2 remedial intervention. The intervention resulted in a faster rate of improvement than that of typical students, and the student reduced the gap. The student “responded to the intervention,” and special education did not need to be considered.

**The Research Agenda: Social Validity and the Role of Rate of Improvement**

This article began with a brief review of the literature on SLD identification, first from a “do as I say” perspective, which considered

![Figure 5. Results of reading curriculum-based measurement (R-CBM) testing comparing three Grade 2 students with different educational needs to local norms in the number of words read correctly. (Copyright © 2007 Edsformation, Inc. Reprinted with permission.)](image-url)
studies that examined whether school-identified SLD students matched a state or school district required criterion, typically the ability–achievement discrepancy. School-identified SLD students often failed to meet the stated eligibility criterion. For more detail, see MacMillan and Siperstein (2002). A brief review was then presented, trying to distill if any consistent pattern in assessment results could be discerned from what schools actually did in the eligibility determination process. I called this the “do as I do” approach. What schools appeared to do when identifying students as SLD was to base this decision on severe low achievement, much lower than typically achieving students, and reliably and significantly lower than students that schools considered low achieving. Further analysis of the features of what schools do in the SLD identification process posited that the achievement performance discrepancy is a relative one, where SLD students’ discrepancies are severe compared to local community peers.

I then illustrated how schools use CBM, a scientifically based progress-monitoring tool, to operationalize a dual-discrepancy approach to RTI, consistent with the recommendations of groups like the National Research Council and leaders in our field. RTI is best conducted within a process of problem solving in a three-tier model where universal screening and progress monitoring are continual activities for all students.

That an RTI process with referred students and embedded within a three-tier model can be done is without question. It has and is being done in schools nationwide. However, actual practice and research-based practice are not always the same. A number of pressing research questions need to be addressed. I have characterized these questions as ones of (a) social validity of specific learning disability, (b) challenges in assessing response to high-quality interventions, (c) technical issues in progress monitoring, (d) fidelity of implementation, and (e) generalizability.

Social Validity of Specific Learning Disability

The identification of a specific learning disability as result of a RTI process with a dual-

![Figure 6. Benchmark assessment of a Grade 2 student showing response to intervention. R-CBM: reading curriculum-based measurement; WRC = words read correctly. (Copyright © 2007 Edformation, Inc. Reprinted with permission.)](image)
discrepancy approach has been and likely will continue to be a topic of intense debate. Many in the field appear to be comfortable with SLD as a solely “within-the-child” phenomenon, even if the associated assessment practices have psychometric deficiencies and/or there is consistent evidence that schools do not use the results as intended.

An RAD approach whereby students labeled SLD are the lowest-achieving students relative to other students in their communities fit the obtained data and observed school practices of the past 25 years. But such a pragmatic conception of a “disability” is uncomfortable to many and threatens their world views. From a social validity perspective, more research is needed on this topic. I see at least three straightforward questions:

1. What threats do professionals (e.g., changing roles) and parents (e.g., loss of disability “uniqueness”) see in an explicit needs-driven service delivery system?
2. What belief systems are challenged when services are weighted toward solving the problem instead of identifying the presumed cause of the problem?
3. How willing are consumers (e.g., school psychologists, teachers, parents, students) to endorse a service delivery system that emphasizes students’ needs (versus labels)

yet ensures the legal protections that some students and families need?

It is entirely plausible to me that we will continue to say one thing (i.e., a student “has” a disability) while doing another (i.e., serving students with severe educational needs relative to community expectations). I would like to understand why.

Challenges in Assessing Response to High-Quality Intervention

To be certain, it has been evidenced that historically, a severe RAD has been the defining feature of students whom schools label as SLD. This discrepancy is easy to measure and is important to teachers and parents. RTI is based on dual discrepancy, however, and it is a discrepancy in response to high-quality intervention that presents two obvious challenges that must be addressed.

“Response” to date appears to be assessed best by calculating a rate of improvement (ROI) in a fixed period of time using CBM. The first RTI challenge is schools’ limited capacity to measure response or ROI. Despite more than 25 years of research and practice with CBM as a frequent progress-monitoring tool, it remains an underutilized tool in schools. It even remains underutilized for the purpose and the population of students for whom it was originally intended, monitoring of progress toward IEP goals for students who receive special education. Without CBM, accurately assessing response to intervention will be difficult.

The problem of limited CBM use as a progress-monitoring tool is not with school psychologists. Shapiro, Angello, and Eckert (2004) report that more than 90% of school psychology training programs instruct their students in its use and more than half of practitioners include CBM among their assessment tools. However, school psychologists are not the targeted audience for CBM progress-monitoring skills. We must understand the limited amount of progress-monitoring training at the preservice and in-service levels for teachers and administrators. What are the barriers to a wide range of teachers receiving scientifically based progress monitoring training? How do
schools ensure all their teachers have the capacity for assessing all students’ response to intervention accurately and efficiently, including those considered for SLD eligibility?

The second RTI challenge is ensuring high-quality intervention. It is beyond the scope of this article to identify all the research needs here. However, we must continue to try to understand what constitutes high-quality interventions in educators’ knowledge base and skill sets, as well as what barriers there may be to their use.

Because of these two challenges (i.e., limited CBM progress-monitoring skills, implementation of high-quality intervention), it is possible that this part of the dual discrepancy will be ignored in practice. The research agenda should be designed to include investigating how much attention and weight decision makers give to a student’s ROI using scientifically based progress monitoring (e.g., CBM) above and beyond the achievement discrepancy. This question can be answered in an objective manner through examining the explained variance in placement decisions contributed by the achievement discrepancy and ROI. Theoretically, ROI should add additional explained variance in the eligibility decision. Alternately, RTI individual case studies should be examined through systematic analyses of achievement discrepancy and ROI decision criterion. Although intellectually and practically appealing, it will be interesting to see if both discrepancies contribute to the decisions school make. There is precedent for this line of research in judging SLD students’ response to reintegration in general education after being served in special education resource rooms (Shinn, Powell-Smith, Good, & Baker, 1997).

**Technical Issues in Progress Monitoring**

Although there are a number of what I consider “macro issues” (e.g., social validity of disability) in RTI, several research issues relate to the specific technical issues in progress monitoring. Three issues that are paramount are goal setting, sensitivity to improvement, and reliability of judgments of adequate progress. These issues, of course, are interrelated. RTI goal-setting practices are short-term variations of long-term goal-setting strategies (e.g., annual IEP goals) such as growth rate standards (Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993) or “reducing-the-gap” normative approaches (Shinn, 2005a). For annual goals, a goal is set for a 25- to 32-week period, students are assessed one to two times per week, and judgments of actual progress are made based on a trend line drawn through a sufficient number of accumulated data points, preferably at least 7–10, compared to the expected rate of progress (Shinn et al., 1989). Notably, the decision about progress is relatively “low stakes”: whether the student is progressing and whether the intervention needs modification at a given point in time. Judgment errors are, in a sense, self-correcting as additional progress data are collected and the need for change in intervention may be more apparent.

In a RTI process, goals are typically set for 6–12 weeks and students are assessed one to two times per week. This limited time frame makes it potentially challenging to accumulate a sufficient number of data points for reliable judgments. With RTI, the decision about progress is also more high stakes and less self-correcting. Failure to respond potentially results in a special education label and placement where a more intensive intervention may be implemented that may be accompanied by well-known special education side effects such as stigmatization, reduced time with non-disabled peers, and so forth.

We would expect that long-term goal-setting and progress-monitoring strategies would result in a generalizable process to RTI. However, we need to generate much more information regarding the nuts and bolts of the progress-monitoring practices that work best to produce quality decisions that result in special education. Among the important research topics are frequency of measurement, number of data points, goal ambitiousness, and reliability of the progress-monitoring judgments made, including those made by parents as well as educators. For example, in RTI goal setting, one of the methods is based on reducing the achievement gap from typically achieving students. How much of a reduced gap is seen as
socially important? Would there be the same consensus regarding social importance among educators as with parents? Although we know that with appropriate training reliable judgments about progress can be made by educators (e.g., Shinn et al., 1997), we know very little about how parents reach conclusions about progress. More studies on how professionals and parents derive their goals and reach progress decisions are in great need.

Fidelity Issues

As with all school implementation issues, I look forward to empirical data on how accurately schools engage in a RTI process when progress-monitoring and intervention fidelity data are to be collected and used. Time-efficient assessment is no guarantee that the data are collected and used with integrity. Fidelity of treatment implementation is critical not only as part of the research agenda, but also to ensure the integrity of the decisions made with students.

Generalizability

To date, nearly all published illustrations of RTI have been in reading. Some students also have deficits in other language arts areas (e.g., written expression, spelling) and in mathematics computation and applications. Whether the same assessment strategies for determining the severity of achievement discrepancies, the types of data collected, goal setting, and frequency of progress monitoring are applicable is an open question. Certainly there are quality progress-monitoring tools in these areas, but compared to reading, they are significantly underinvestigated in general and seriously so in the RTI process.

Finally, from a legal perspective, RTI is an eligibility option for local education agencies only for learning disabilities. Historically, the concepts of the dual discrepancy (i.e., educational need and quality and intensity of intervention required for educational benefit) fleshed out explicitly in RTI for SLD are implicit for all students with severe educational and behavioral needs in a problem-solving model. A fundamental question becomes whether the field of education can go further and apply these fundamental concepts to all students regardless of label. I look forward to seeing how the field addresses all these challenges and more in our attempt to improve outcomes for all students.

References


Mark R. Shinn, PhD, received his doctoral degree from the University of Minnesota in 1981. He currently is Professor of School Psychology at National Louis University in Skokie, Illinois, and serves as Project Director for IASPIRE North, a U.S. Department of Education and Illinois Department of Education Personnel Preparation Grant. IASPIRE was designed to support schools’ implementation of a three-tier problem-solving model, including RTI. A recipient of the 2003 American Psychological Association’s Jack Bardon award, he continues his research and training efforts to support practitioners to use scientifically based progress-monitoring tools.