INTRODUCTION

PART I

History of Learning Disabilities and Emergence of a New Model

Chapter 1

LEARNING DISABILITIES: DEFINITION AND BACKGROUND

The concept of learning disabilities dates back to the early 1960s. In 1968 the label of "specific learning disability" was added as a federally designated category of handicapping conditions (Hallahan, Kauffman, & Lloyd, 1999). One of the first to address the definition of learning disabilities was Samuel Kirk. In 1962 Kirk wrote:

> A learning disability refers to a retardation, disorder, or delayed development in one or more of the processes of speech, language, reading, writing arithmetic, or other school subject resulting from a psychological handicap caused by a possible cerebral dysfunction and/or emotional or behavioral disturbances. It is not the result of mental retardation, sensory deprivation, or cultural and instructional factors (Kirk, 1962, p. 263).

In Kirk's description can be seen many components of the modern definition including a conceptualization that LD (1) is a deficit in processing (2) that results in reduced academic performance in one or more areas, (3) is possibly related to a cerebral (pertaining to the central nervous system) dysfunction, and (4) is not the result of other handicapping conditions. Later in 1965, Barbara Bateman proposed a modified definition of learning disabilities that removed emotional factors as causal in LD and more significantly suggested that it could be identified by an "educationally significant discrepancy" between estimates of intellectual potential and actualperformance level (for discussion, see Hallahan, Kauffman, &

Lloyd, 1999; Smith, 1998). This discrepancy notion was further supported by the epidemiological work of Rutter and Yule in the early to mid-1970s. By studying the IQ predicted reading achievement of children ages 9 to 13 on the Isle of Wright they concluded that there was an abnormal distribution of reading performance scores suggesting that (1) reading underachievement occurred at a higher than expected rate and (2) that different patterns of sex distribution and of neurological deficit and development were observed in the "under achievement" group (Rutter & Yule, 1975). Thus support for the first severe discrepancy provisions for learning disabilities emerged.

THE HISTORY OF LD

Arguably the most important landmark legislation providing rights and educational privilege to students with disabilities was PL 94–142 enacted by Congress in 1975. Prior to 1975 approximately 200,000 individuals with significant disabilities were institutionalized in state-run settings and generally provided minimal standards of care (Ed.gov. 5/21/2007). Further, in 1970 only one in five children with disabilities was educated in public schools. Perhaps one of the most debated classification categories in the PL 94–142 regulations was with respect to learning disabilities.

While crafting a definition of LD in 1976 for the PL 94–142 regulations, the United States Department of Education (USDOE) considered the addition of a severe discrepancy formula (e.g., achievement falling 50 percent or more below the child's expected achievement level) within the LD definition. While these efforts were offset by a number of objections from national experts of the time offering an array of conceptual and statistical difficulties with this procedure, the notion of seemingly objective discrepancy criteria was not entirely abandoned. The final definition for learning disabilities in PL 94–142 was as follows:

> The term "specific learning disability" means a disorder in one or more of the basic psychological processes involved in understanding or using language, spoken or written, which may manifest itself in an imperfect ability to listen, speak, read, write, spell, or to do mathematical calculations. The

> term includes such conditions as perceptual handicaps, brain injury, minimal brain dysfunction, dysfunction, dyslexia, and developmental aphasia. The term does not include children who have learning disabilities which are primarily the result of visual, hearing, or motor handicaps, or mental retardation, or emotional disturbance, or of environmental, cultural, or economic disadvantage. (U.S. Office of Education, 1977, p. 65083)

While the actual definition in the pivotal regulations did not include a severe discrepancy formula, the section of the law that identified criteria for identifying students with learning disabilities stipulated that:

a. A team may determine that a child has a specific learning disability if:
   1. The child does not achieve commensurate with his or her age and ability levels in one or more of the areas listed in paragraph (a) (2) of this section, when provided with learning experiences appropriate with the child's age and ability levels; and
   2. The team finds that a child has a severe discrepancy between achievement and intellectual ability in one or more of the following areas:
      i. Oral expression;
      ii. Listening comprehension;
      iii. Written expression;
      iv. Basic reading skill;
      v. Reading comprehension;
      vi. Mathematics calculation; or
      vii. Mathematics reasoning
b. The team may not identify a child as having a specific learning disability if the severe discrepancy between ability and achievement is primarily the result of:
   1. A visual, hearing, or motor handicap;
   2. Mental retardation;
   3. Emotional disturbance; or
   4. Environmental, cultural, or economic disadvantage. (Federal Register, Dec. 29, 1977, p. 65083)

Therefore, while the severe discrepancy language did not make it into the formal LD definition, the inclusion of the preceding language essentially added these procedures to the classification. Following the publication of PL 94–142 most states adopted severe discrepancy provisions in their identification procedures for learning disabilities (e.g., Frankenberger & Franzalglio, 1991). However states varied in terms of the tests used to ascertain a discrepancy, the formulas used to compute the discrepancy, and the magnitude required for identification purposes (for discussion, see Fuchs, Mock, Morgan, & Young, 2003).

CRITICISMS OF DISCREPANCY-BASED MODELS

Criticisms of discrepancy-based models for understanding and identifying learning disabilities are numerous and have a long history. Essentially these criticisms can be conceptualized along two domains: problems with the reliability of a discrepancy-based approach for identifying students with disabilities and problems with the discrepancy-based model for conceptualizing and treating students with learning disabilities. Therefore, basic criticisms of the discrepancy-based model are that this method for understanding and identifying learning disabilities lacks adequate reliability and validity. In terms of reliability, the 300 percent increase noted in the population of students identified with learning disabilities over the last 30 years suggests a lack of stringent criteria for making the diagnosis (President's Commission on Excellence in Special Education, 2002).

PROBLEMS WITH RELIABILITY

One specific difficulty hampering reliable diagnosis is that there are four major methods for determining the presence of a severe discrepancy and each uses different criteria. The methods include assessing the discrepancy in terms of (1) deviation from grade level, (2) Standard deviation from the mean, (3) Standard Score comparison, and (4) Standard Regression analysis. The first, *deviation from grade level*, suggests that if Kate is in the fourth grade yet reads at a second-grade level then she may be seen as having a severe discrepancy in her reading achievement. In this method

Kate's academic performance is compared to her peers. The second method, *standard deviation from the mean*, might assess Kate on an individually administered achievement test. Given that her score overall or in a specific academic area was at least a standard deviation below the norm she may be perceived as evidencing a severe discrepancy commensurate with an LD diagnosis. This method would compare Kate's achievement with that of a standardized sample of same-age students from across the country. In the third method, *Standard Score comparison*, Kate's performance on an individually administered intelligence test would be compared to her performance on an individually administered achievement test. If she achieved an IQ score of 100 (average score) and an achievement score one or more standard deviations below the mean, she may be seen as evidencing a severe discrepancy commensurate with an LD diagnosis. With this method Kate's academic performance is compared to her performance on an intellectual assessment. Given that the comparison groups for Kate's academic performance differ across these three methods (e.g., compared to peers, a national sample, and to her own IQ score), it is not hard to imagine why the result would be different for students diagnosed as learning disabled depending on the discrepancy method utilized. In essence, different methods of calculating a discrepancy will result in different students being classified. The fourth method, *Standard Regression analysis*, utilizes the Standard Score comparison technique and additionally employs a regression formula as an attempt to statistically account for the measurement error associated with the tests, the reliability of them, and the correlations between them. While this is perhaps the most psychometrically sound method for assessing IQ/achievement discrepancies, it is not without additional inherent difficulties.

In a replication of an earlier study Mercer, Jordan, Allsopp, and Mercer (1996) surveyed all state education departments in the United States and found that 98 percent of them included a discrepancy in their definition of and identification criteria for learning disabilities. As indicated in the 1997 NYS Part 200 Regulations of the Commission of Education, "a student who exhibits a discrepancy of 50% or more between expected achievement and actual achievement determined on an individual basis

shall be deemed to have a learning disability." This determination in contemporary assessment was often completed using an intelligence test as the measure of *expected achievement* and a norm-referenced, standardized, academic test as a measure of *actual achievement*. The difference between the two scores is used to assess the discrepancy.

This brings us to the second major difficulty significantly hampering the reliability of LD diagnoses made with discrepancy based methods: The norm-referenced, standardized measures commonly employed in this assessment process are inadequate for measuring both *expected achievement* and *actual achievement*. In terms of expected achievement, while IQ tests are good general predictors of educational attainment they are inadequate for assigning an expected achievement outcome for individual students for several reasons. First, IQ test components most linked with reading performance are often verbally mediated and are somewhat dependent on reading. Therefore poor readers may have lower verbal IQ test scores and therefore be denied special education services due to a lack of assessed discrepancy (see Siegel, 1989; Stanovich, 1989). Secondly, this approach assumes that IQ can accurately predict academic performance. To explore this further we can look at the correlations between IQ and achievement reported on the most recent version of a popular standardized achievement measure, the Wechsler Individual Achievement Test-Second Edition (WIAT-II, 2002). The examiner's manual of the WIAT-II reports that the correlations between full-scale ability (assessed by the WISC-III) and achievement (assessed by the WIAT-II) range from .3 to .78. To understand how well the WISC-III predicts achievement we can square these correlations to determine the amount of shared variance between these scores. The result suggests that the WISC-III accounts for 9 to 61 percent of the variance in a given student's achievement test score. This also suggests that from 39 to 91 percent of the student's achievement score is not accounted for by the IQ test. This lends considerable doubt to the notion that an IQ test can accurately assign an expected level of achievement, at least at the level of the individual student. Second, with respect to *actual achievement*, the concept that a student's actual academic performance can best be assessed with a norm-referenced test

administered at a single point in time has received considerable criticism as well. Among these criticisms are that nationally normed standardized achievement assessments often do not reflect the skills in a given local curriculum, they suffer from regression to the mean effect, and the fact that all psychometric tests include measurement errors that vary across students and across characteristics of the student (see Francis, Fletcher, & Morris, 2003). In single point assessments measurement error creates fluctuations in test scores that vary by test, age, ability level, and ethnicity. Applying cut-off scores to these types of score distributions is problematic since there is generally little or no actual difference between children at or around that cut-off regardless of their assigned status. Score fluctuations (above or below assigned cut-off scores) have been assessed in both real and simulated data sets suggesting that up to 35 percent of cases change status based on measurement error when single tests were used. Similarly, with respect to discrepancy scores, actual data from the Connecticut Longitudinal Study, analyzed by Francis et al. 2005, found that approximately 20 percent to 30 percent of students studied change disability status from third to fifth grade based on discrepancy scores.

Given the cited limitations with the discrepancy model it is easy to see how it lacks reliability in diagnosis. The fact that different criterion are used across different states significantly impairs consistency in identification. In addition, the limited ability of IQ tests to predict the achievement of an individual measurement error, and the difficulties associated with assigning cut-offs in either single test or discrepancies between tests significantly limit the reliability of this approach. In sum the use of discrepancy-based psychometrically oriented models for diagnosis are unreliable and insufficient to accurately designate individuals with learning disabilities (Francis, et al., 2005; Fletcher et al., 2005).

Problems with Validity

In addition to reliability concerns, discrepancy-based models also have been heavily criticized with respect to validity. Since the validity of a construct relies on its uniqueness and utility, the validity of the discrepancy-based model assumes that IQ-achievement

discrepant students are qualitatively different from "regular" (non-discrepant) low achievers. If this model were valid, these two groups of students would differ in terms of their prognosis (development of reading ability), response to intervention (discrepant and non-discrepant groups should show differential response to reading intervention), and with respect to the cognitive profiles thought to underlie reading abilities (e.g., Francis et al., 1995).

The literature in this area has been generally unsupportive of the discrepancy-based model for LD classification. Studies by Stanovich and Seigel (1994) and by Fletcher et al. (1994) suggest that IQ discrepant and non-discrepant low-achieving groups did not differ on measures of independent reading ability. The two groups were also found to have no significant differences with respect to cognitive abilities believed to underlie reading development. Both of these independent studies found that language-based measures were better predictors of early reading ability than performance on IQ tests. In addition several meta-analyses have found little difference in the cognitive process of IQ discrepant and non-discrepant low-achieving groups and further that these groups did not differ with respect to reading development (e.g., Hoskyn & Swanson, 2000; Stuebing et al., 2002).

With respect to Response to Intervention (RTI), a series of studies conducted by Vellutino and colleagues at the University of Albany provided longitudinal data on students' reading development. In one study incorporating an intense reading remediation component, Vellutino, Scanlon, and Lyon (2000) followed 118 impaired readers and 65 control students from kindergarten through third grade. Their findings suggested that IQ scores could not distinguish between impaired and normal readers, nor were IQ scores helpful in predicting impaired readers who were difficult to remediate versus impaired readers who were readily remediated. Further, they found that in normally developing students IQ did not predict reading achievement nor was it correlated highly with measures of reading ability (e.g., word identification, phonological decoding). The conclusion of this study was that "when intense remediation resources are made available to impaired readers representing a broad range on the intellectual continuum, response to remediation is not associated with measured intelligence" (p. 237).

In sum the literature investigating the validity of discrepancy models has generally been unsupportive. Findings suggest that IQ discrepant and IQ non-discrepant groups of students do not differ in terms of their cognitive profiles, their prognosis in reading, or in their RTI. Included in the President's Commission on Excellence in Special Education Report (USDOE, 2002) are two poignant quotes included here, the first from Dr. Sharon Vaughn and the second from Commissioner Wade Horn.

> *"There is no compelling reason to use IQ tests in the identification of learning disabilities. And, if we eliminated IQ tests from the identification of individuals with learning disabilities, we could shift our focus on making sure that individuals are getting the services that they need and away from the energy that's going into eligibility determination." (p. 22)*
>
> *"I would like to encourage this Commission to drive a stake through the heart of this over reliance on the discrepancy model for determining the kinds of children that need services. It doesn't make any sense to me. I've wondered for 25 years why it is that we continue to use it and over rely on it as a way of determining what children are eligible for services in special education." (p. 25)*

A PLACE FOR INTELLIGENCE TESTING IN LD DIAGNOSIS?

Criticisms aside, there are reasonable proponents for the continued use of intelligence testing in the assessment process for learning disabilities. Fuchs, Mock, Morgan, and Young (2003) aptly point out that the empirical support for the relationship between IQ and school achievement has a history spanning more than 50 years. Further they illustrate the practical pedagogical implications associated with instruction provided to all children ranging in IQ from 70 to 155. The authors contend that IQ testing may help to preserve our historical conceptions of LD as a distinct diagnostic category understood as children failing to learn with average or above-average intelligence and adequate participation in the general curriculum. In their emerging model for LD assessment the authors suggest that children not responding to an effective general education curriculum, and demonstrating a lack of response to a more intense level of instructional

intervention, would then be administered "valid cognitive assessments" to "facilitate identification of students with LD in historical terms." There does seem to be some utility in administering cognitive assessments to students demonstrating a failure to progress despite increasingly intense instructional interventions shown to be effective for others. This utility may be ruling out cognitive delays as potentially impairing academic functioning and in providing additional information relevant to the diagnosis. Based on the literature previously provided however, it would appear that the routine use of IQ tests to ascertain an aptitude-achievement discrepancy with a single point-in-time assessment may not be warranted.

EMERGENCE OF CONTEMPORARY MODELS OF LD

HISTORY

The period from the late 1970s to the mid 1990s evidenced a substantial increase in the numbers of students identified with specific learning disabilities (SLD). Summaries of prevalence data over this timeframe suggest that SLD rates have risen by as much as 300 percent, that roughly 80 percent of these SLD students evidenced unaddressed deficits in reading, and that as many as 40 percent to 50 percent of all children served in special education had not been adequately instructed in reading (i.e., President's Commission on Excellence in Special Education, 2002; U.S. Office of Special Education, NJCLD, 2002).

In response to the apparent overidentification of students as learning disabled, prereferral intervention and the use of prereferral intervention teams became popular by the mid to late 1980s. These teams typically provided for collaborative consultation to teachers toward instructional modifications or accommodations to increase student performance prior to or without special education referral. Perhaps the earliest widespread teaming initiative was the Teacher Assistance Teams reported on by Chalfant, Pysh, and Moultrie (1979). These teams were formed to provide an avenue for teachers to assist other teachers in intervening with and accommodating for students seen as difficult-to-teach. While some data was reported in terms of large-scale implementations of

these teams, by in large they were not well studied nor were data on student progress or the integrity of the team process typically provided (e.g., Fuchs, Mock, Morgan, & Young, 2003).

As prereferral intervention teams evolved, many adopted components of behavioral consultation that structured the problem-solving process and relied on student monitoring data to inform intervention design and revision. A major facilitator to this evolution was the emerging literature supporting the use of Curriculum-Based Measurement (CBM) procedures and behavioral observation methods to monitor student's academic and behavioral progress in the schools. Research has generally indicated that team consultation procedures and quality prereferral interventions can lead to increased student performance and decreases in special education (Graden, Casey, & Bonstrom, 1985; Flugum & Reschly, 1994; Rosenfield, 1992).

Perhaps the most comprehensive evaluation of prereferral intervention teams was conducted by Doug and Lynn Fuchs and colleagues. These researchers, supported by the Office of Special Education Programs, developed and evaluated a Mainstream Assistance Team (MAT) model implemented in the Metro-Nashville school district in the mid to late 1980s. The Fuchs utilized graduate students, a scripted behavioral consultation process, and prescriptive interventions to address teacher concerns of referred students. Over the three-year project the Fuchs demonstrated that these procedures lead to a significant decrease in special education referrals, 75 percent or more of referred students meeting teacher generated goals, and high teacher perception of effectiveness. While effective, the MAT project did not sustain much past the life of the grant funding (Fuchs et al., 1996), due perhaps to the prescriptive nature of the program or the artificial/external development and infusion of the project into the host site.

Based on the work of the Fuchs the School Based Intervention Team (SBIT) model was developed and implemented in a large urban district in Central New York (McDougal, Clonan, & Martens, 2000). This team-based prereferral intervention model was similar to MAT's in that it followed the behavioral consultation process but was developed and implemented largely by in-district personnel. The authors' contention was that by creating an "in-house" model and by attending to principles of organizational

change an effective and sustained project would result. A two-year evaluation of four SBIT teams indicated that they were successful in decreasing special education referrals by 40 percent compared to similar non-SBIT schools and that referred students overall evidenced significant increases in academic and/or behavioral performance (McDougal et al., 2000). In terms of sustainability, while SBITs still function in the host district after the initial report, the functioning and effectiveness of these teams has not since been reported.

Although considerable research has documented the effectiveness of high-quality prereferral intervention, the process has some inherent structural limitations. Prereferral intervention teams generally provide intervention services toward individual students referred by the classroom teacher. This requires (1) that the student's level of functioning is poor enough to generate considerable teacher concern, (2) that the classroom teacher is willing to take the time required to complete the referral and intervention team process, and (3) that the team itself has adequate time to devote to the referral. In essence, even good models for prereferral intervention suffer some of the inherent limitations found in traditional approaches for special education identification, including an approach that waits for students to evidence failure; idiosyncratic student identification based on teacher referral; a focus on individual students rather than whole classrooms, grade levels, or schools; and the development of individualized interventions as opposed to systemic development of increasingly intense resources based on student need.

RTI: A General Definition

Given the preceding concerns more systemic large scale implementations of Response-to-Intervention models have been initiated in the schools. In reviewing the literature, Fuchs et al. (2003) identified critical components of RTI implementation. In general Fuchs et al. suggest that:

"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" (p. 159).

Based on this general description four large-scale implementations have been identified and reported in the literature. Two large-scale implementations of RTI include Ohio's Intervention Based Assessment (IBA) and Pennsylvania's Instructional Support Teams (ISTs). These implementations focused on the provision of prereferral intervention services prior to the referral to special education. The other two models of implementation, Heartland Agency (Iowa) and Minneapolis Public School's Problem-Solving Model (MPSM) actually utilize student progress monitoring data to designate eligibility for a "non-categorical" designation into special education. Each of these large-scale RTI implementations has their roots in the team-based prereferral intervention model incorporating a collaborative team approach following a behavioral consultation model. Each model is briefly described below.

The First Large-Scale Implementations of RTI

The Ohio IBA and the Pennsylvania IST models were similar in that they were both large-scale implementations of structured prereferral intervention team initiatives. The IBA started in 1992–93 as a volunteer initiative supported by a special education waiver from the Ohio State Department of Education. In the beginning 35 schools from across the state were recruited to provide school-based intervention services to struggling students. They followed a behavioral problem-solving process including problem definition, collection of baseline data, setting specific goals, hypothesizing a reason for the student difficulty, developing an intervention and monitoring plan, and evaluating the intervention via student progress data compared to baseline levels (e.g., Telzrow, McNamara, & Hollinger, 2000). An evaluation of IBAs reported by Telzrow in 2000 indicated that by 1997, 329

school teams were running though only 227 were included in the study. In reviewing "best case" documentation from participating teams, Telzrow and colleagues found that most teams were not applying the problem-solving model with integrity especially as related to documenting that the developed student interventions had been implemented as designed (for discussion see Fuchs, Mock, Morgan, & Young, 2003).

The Pennsylvania IST model was also supported by the respective State Department of Education. In addition to the collaborative and behavioral problem-solving process, the IST model utilized a full-time support teacher responsible for assisting the classroom teacher to implement student interventions developed by the team. This student-directed support is monitored continuously with CBM and/or behavioral assessment and periodically evaluated to refine intervention procedures. The IST support is limited to 50 school days when the team meets to decide if further evaluation is warranted. Evaluations of the IST model suggest that teams successfully follow the prescribed model, that ISTs lead to decreases in special education referrals and placements, reductions in the use of grade retentions, and increases in general measures of student behavior and academic performance (e.g., Hartman & Fay, 1996; Kovaleski, Gickling, Morrow, & Swank, 1999). Overall while these results seem promising the IST evaluations have been criticized for a lack of direct measures both of team functioning and student outcomes, a lack of inter-rater reliability data, and a lack of specific descriptions of the interventions utilized (Fuchs et al., 2003).

Perhaps the pioneer in using RTI procedures for eligibility determinations is Heartland, Iowa's largest educational agency providing technical assistance and training to 350 schools across 56 districts. Heartland's reform dates back to the mid 1980s when state support for noncategorical models of special education and direct assessments of student performance emerged. The Heartland model utilized a four-level (tier) model for intervention and assessment. The levels include (1) collaboration between the teacher and the parent, (2) referral to the Building Intervention Team, (3) referral to district staff (e.g., school psychologists and special educators), and (4) considerations of special education. In the Heartland model student academic level and learning rate are

compared to local classroom or grade-level norms to ascertain the need for increasing levels of support. This model is noncategorical in that low-performing students are not ultimately labeled as LD, MR, and so on, but rather as students eligible for special education services. Again, evaluations of the Heartland model suggested generally positive results in terms of reductions in special education referrals and placements and with regard to student performance, though these studies too were critiqued for a lack of empirical rigor (i.e., Ikeda, & Gustafson, 2002; Fuchs et al., 2003).

The Minneapolis Public School's Problem-Solving Model (MPSM) was developed in the early 1990s as an assessment and intervention model to in part reduce the overrepresentation of minorities in special education and also to create a focus on instruction and student performance. Similar to the Heartland model, MPSM began as a four-tier behavioral problem-solving process. As with the Heartland model, student academic level and rate of progress as compared to local norms is used to assign the need for increasingly more intense levels of service. In addition, MPSM too uses lack of student progress to decide eligibility for special education placement and also employs a noncategorical approach to identification. Published evaluations of MPSM have suggested more proportionate representation of minorities in special education, a stable overall identification rate of approximately 7 percent, and increases in referred students' academic performance especially in the area of reading (Marston, Muyskens, Lau, & Canter, 2003). These studies too were critiqued with respect to the empirical rigor of the employed measures and design (Fuchs et al., 2003).

SUMMARY

In reviewing the major field-based implementations of RTI, all utilized a collaborative team-based approach and followed a behavioral model of consultation to format the problem-solving process. The IBA and IST initiatives were large-scale, state-supported models for providing prereferral interventions to students prior to (or instead of) referral for special education eligibility determination. The Heartland and MPSM projects

were implemented in large district (or "agency") settings, and while similar to IBA and ISTs, these two models utilized student progress data compared to class or grade level norms to ascertain eligibility for special education. In addition both Heartland and MPSM employed a noncategorical approach to special education, not requiring school professionals to conduct further diagnostic assessment to assign the student to a diagnostic category. In general, data from these implementations are promising though far from complete. While reductions in special education referrals and placements are noted along with increases in students' academic and behavioral performance, most of these studies suffer methodological difficulties inherent in field-based evaluations. These difficulties include reporting on incomplete and restricted samples, lack of direct performance or integrity measures, lack of detailed intervention protocols, use of "convenience" data sets, and a lack of consistency in measures/outcomes employed across studies. Further, while not discussed here, prior reports of these implementations also include considerable concerns with the amount of professional development required for implementation and the difficulty maintaining consistency in model implementation both across time and school setting.

RESEARCH-BASED RTI MODELS

In addition to field implementations several research-based RTI models have been implemented. These research models typically employ a standard protocol approach to intervention as opposed to a problem-solving approach. This standard protocol approach offers the same empirically based treatment to all children identified with low skills. The advantages to a standard protocol procedure is that it is easier to validate, train practitioners, and measure the integrity of one intervention being implemented as opposed to many possibilities derived from the problem-solving models (e.g., Fuchs et al., 2003).

Research-based models have generally been implemented in the primary grades with struggling readers. Vellutino et al. (1996) tracked the literacy development of a significant sample ($n = 1407$) of children from kindergarten through fourth grade. In the study intense reading intervention was provided to low readers

from the middle of first grade to the middle of second grade (depending on student progress). Interventions were daily 30-minute tutoring sessions focusing on phonemic awareness, decoding, sight words, comprehension, and text reading. Students were provided 70–80 sessions over the course of a semester. Based on this work Vellutino found that two-thirds of tutored readers demonstrated "good" or "very good" rates of progress. Further, based on normative measures, they had reduced the percentage of children classified as disabled readers to 3 percent by the spring of second grade (see Vellutino, Scanlon, Small, & Fanuele, 2003). While employing a rigorous methodology, the study's results have been critiqued as Vellutino excluded students with IQ scores below 90 and due to concerns that schools would not have the resources to replicate these results by providing struggling readers with 70 to 80 one-on-one tutoring sessions with highly trained educators (e.g., Fuchs et al., 2003).

An additional study employing standard protocol procedures focused on second-grade students at risk for learning disabilities. Vaughn, Linan-Thompson, and Hickman (2003) provided three tiers of reading interventions delivered sequentially based on student progress. Low-performing students were initially provided 10 weeks of empirically based supplemental small-group instruction. Based on predefined criteria students were either dismissed from supplemental services based on their progress or grouped for another 10 weeks of instruction (either at the same level of service or a more intense level). After 30 weeks 75 percent of students had been returned to general education and 25 percent referred for special education consideration. Thus they provided an empirical illustration of a preventative three-tier model for RTI: "with primary intervention consisting of the general education program, secondary intervention involving the fixed duration, intensive, standard protocol trial (with the goal of remediating the academic deficit rather than enhancing general education), and tertiary intervention synonymous with special education" (Vaughn & Fuchs, 2003, p. 139). While the results achieved by Vaugh et al. are quite promising, follow-up of discontinued students suggested that a subgroup did not continue to thrive in general education and could have benefited from continued delivery of supplemental services.

SUMMARY

Research-based models of RTI have been successfully implemented in the schools. These models have focused on the use of scientifically based standard protocol interventions provided in the area of literacy to struggling primary grade students. In general these models have demonstrated significant efficacy and the promise to effectively intervene with the majority of referred students without necessitating the need for special education referral. Serious questions remain, however, in terms of the feasibility of these practices for widespread use in the schools based on the training and resources required to implement them. In addition, as the study conducted by Vaughn et al. points out, at least some of these remediated students may require ongoing support in order to progress in general education.

OUR PERSPECTIVE: WHERE ARE WE NOW?

Currently we are at a crossroads in the identification of children with learning disabilities. The traditional road traveled has been soundly criticized for employing identification techniques lacking in reliability and validity and assessment procedures lacking in instructional utility and fraught with cultural bias (e.g., Gresham, 2002). Further, traditional practices employ a "wait to fail" procedure relying on teachers to recognize that a student demonstrates severe learning problems prior to referring them for support services. This imprecise screening practice has lead to students with instructional delays referred to special education, those with learning disabilities not referred, and in general all services being provided too late to remediate children such that they can respond to the general education curriculum (Bradley, Danielson, & Doolittle, 2005; Vaughn & Fuchs, 2003). In response, legislative initiatives and professional groups have advocated the charting of a new course, that of Response-to-Intervention. This model holds the promise of early identification and effective treatment, more comprehensive and non-biased identification practices, increases in student performance and accountability, and decreases in the numbers of students referred for special education.

The road toward RTI is still under construction with incomplete results obtained from both research-based models and field-based implementations. Further, RTI implementations have employed a variety of techniques with problem-solving or standard protocol procedures. In general, however, core features of RTI have been identified and include "(a) high quality, research-based classroom instruction, (b) universal screening, (c) continuous progress monitoring, (d) research-based secondary or tertiary interventions, (e) progress monitoring during interventions, and (f) fidelity measures. Decisions about needed services are based on the quality of student responses to research-based interventions (Bradley, Danielson, & Doolittle, 2005, p. 486)."

Based on these features the remainder of the book offers school-based practitioners and those professions in training a potential map for a road that is currently under construction. Based on our review of the literature, our combined research, and our ongoing experience assisting school settings to implement these procedures, we offer this not as a complete work but rather one that reflects what we know at this time. Our model for RTI employs three tiers of intervention and monitoring of increasing intensity as required to produce student progress. We also address eventual considerations for special education eligibility based on the model developed in Iowa combined with the work of the National Research Center on Learning Disabilities. This eligibility model is articulated in Chapter 8 and includes assessments of discrepancies in student performance and rate of progress as well as the demonstration of instructional need. Further, we make a recommendation for a comprehensive evaluation, including cognitive assessment for students suspected of possessing an educational disability. In addition, we address the considerable task of organizational change that is required from schools and districts attempting to implement and sustain RTI initiatives. In conclusion, we hope to summarize the promise of this approach, the current limitations and unknowns, and what we consider next steps in the development of this model for not only the identification of students with learning disabilities but also for the more successful and effective education of all children.

CASE EXAMPLE: MOVING FROM A TRADITIONAL TO AN RTI MODEL: TBD

Located in the Northeastern United States, Baylor Schools is a suburban school district with a student population of about 6,000 students. The district has literacy and mathematics test scores that are about average when compared with districts throughout the state that have resources and student populations similar to Baylor's. The district has five elementary schools (grades K–4), a middle school (grades 5–6), a junior high school (grades 7–8), and one high school.

While Baylor's overall scores on the state accountability tests were generally average, the percent of students in special education programming was 17 percent, which was higher than the state average of 12 percent. Further, relatively few students transitioned out of special education programming once they entered. A district-level task force studied the problem and found that the largest segment of students in special education were those designated with learning disabilities, most generally involving reading. Many of these students had evidenced academic delays in the primary grades and many had received one or more grade retentions prior to classification. The task force also reported that the five elementary schools implemented different core curriculums in literacy and math and had no consistent measures for monitoring students' progress in basic skills other than the state-mandated assessments.

When the task force issued its report the school superintendent, Ann Douglas, had served in her current position for three years. She was knowledgeable about RTI and wanted to establish as a priority the goal of instituting an RTI process in her school district. Referencing the report of the district-level task force, she made a presentation to the school board regarding RTI, highlighting the strengths of this preventative approach in terms of consistent assessment and instruction, screening and progress monitoring, and efficacy in early intervention. Based on this presentation, the Board of Education at Baylor convened a study group to more thoroughly investigate RTI and, if it appeared promising, to make recommendations for the district to move forward with implementation.