One

OVERVIEW OF RESPONSE TO INTERVENTION

For decades the role of educational assessment in the United States has contradicted the very basis upon which education in this country was founded. Data collected by school psychologists and educational diagnosticians for the past 50 years were used to classify students as extremely high and low in order to rank them (Reschly, 1996). As Reynolds (1975) stated, "The dominant orientation in measurements was to a simple kind of prediction that supported the selection of high and rejection of low achievers" (p. 5). However, as early as 1749, Benjamin Franklin wrote in the *Proposals relating to the education of youth in Pennsylvania* that "*all* should be taught to write a fair hand, and swift, as that is useful to all" (Cutler, 1905, p. 56, emphasis added), and the founders clearly saw education as a means to ensure that all citizens could participate in business, express ideas, and fully involve themselves in a democracy (Rothstein & Jacobsen, 2006).

More recently, the Goals 2000 (1994) and No Child Left Behind (2001) legislations continued the line of federal regulations that emphasized the need for all students in this country to be proficient in the basic skills, and the dominant paradigm simultaneously changed from assessment *of* learning to assessment *for* learning (Stiggins, 2005). Assessment in the 1970s and 1980s focused on identifying aptitudes and cognitive processes that were linked to particular disabilities and to learning profiles that could be used to modify instruction. However, decades of research did not support that instructional modifications based on aptitude data led to improved or more robust student learning (Kavale & Forness, 1999). Thus, the U.S. Department of Education's Office of Special Education Programs recommended that measures of aptitude and cognitive processing *not* be used when identifying a child with a specific learning disability (SLD) asserting there is "no current evidence that such assessments are necessary

or sufficient for identifying SLD" (*Federal Register*, 2006, p. 46651). Instead, school districts are now allowed to use a process to determine if a child responds to research-based interventions as part of the SLD identification evaluation. This process, commonly referred to as response to intervention (RtI), is quickly being adopted by school districts all across the country.

There are RtI implementation sites in all 50 states, but what constitutes RtI can be a matter of some debate. In education, there is a long history of widely adopting an innovation without first evaluating its research base or ensuring consistent implementation. When this happens, the innovation that was once hailed as the newest best practice often ends even more abruptly than it began, and "today's flagship" becomes "tomorrow's abandoned shipwreck" (Ellis, 2005, p. 200).

Because a solid research base and consistent implementation are both necessary components of an effective educational innovation (Ellis, 2005), this book will provide the details of both pertaining to RtI. Detailed in the pages that follow is a critique of old models of SLD diagnosis and a summary of the research base for RtI. Chapter 2 provides specific implementation guidelines. RtI is primarily the use of assessment data to make instructional and resource allocation decisions (Batsche et al., 2005; Burns & VanDerHeyden, 2006; Tilly 2008), one of which may be whether to identify a child as having a SLD. Thus, RtI is an assessment process with diagnostic implications. The purpose of this book is to specify characteristics of technically adequate RtI implementations such that accurate diagnostic decisions may be based on the results. A technical adequacy model is necessary to ensure that implementations contain the features that will result in desired implementation outcomes and to aid in the evaluation of implementation efforts in research and practice.

EMPIRICAL ROOTS OF RTI

Problems with the Old System

As stated, the federal provision for RtI came out of special education regulations. Special education was defined by the most recent Individuals with Disabilities Education Improvement Act (2004), as "specially designed instruction, at no cost to the parents or guardians, to meet the unique needs of a child with a disability" (§§ 300.39). Thus, special education relies on two facets: (a) providing effective instruction that is individualized to student needs, and (b) the valid identification of student disabilities. These two facets are highly related because valid diagnostic paradigms are those in which the data lead to interventions with

known outcomes (Cromwell, Blashfield, & Strauss, 1975; Hayes, Nelson, & Jarrett, 1987; Messick, 1995). In other words, the diagnosis should lead to treatments with predictably positive outcomes, and failure to link the two results in a diagnostic framework that is fraught with invalid decisions. As discussed later, special education has a history of difficulties with both of its basic facets. Because RtI currently is allowed in federal SLD regulations, we focus our discussion on SLD.

Specially Designed Instruction

The first aspect of an effective approach to special education is individually designed instruction to provide educational benefit to individual students (*Hendrick Hudson School Board of Education v. Rowley*, 1982). However, previous research found no differences in the instruction delivered to students in special education classrooms as compared to students with and without disabilities in general education courses, or among students in the same special education class (Ysseldyke, O'Sullivan, Thurlow, & Christenson, 1989). Moreover, Glass's (1983) seminal meta-analysis found negative effects for academic and social outcomes for children in special education and concluded that "special placements continue to be made for reasons other than benefits to pupils" (p. 69).

DON'T FORGET

The goal of RtI is to enhance learning for all students including those who are at risk but not identified with a disability (Burns & VanDerHeyden, 2006).

Concern about the effectiveness of special education was certainly an impetus to the RtI provision, but the goal of RtI is to enhance learning for all students including those who are at risk but not identified with a disability (Burns & VanDerHeyden, 2006). Special education used to operate very much like a bounty hunter system where increasing eligibility rates brought more money to a school. However, more recent iterations of the special education mandate allowed for funding to be distributed based on total student population in an attempt to sever the tie between categorized disabilities and monetary contingencies. Recent research has also suggested the need for reform in general education. Generally speaking, less than one-third of students in the elementary grades scored within a proficient range on recent assessments from the National Assessment of Educational Progress in math (Manzo & Galley, 2003), reading (National Center for Educational Statistics 2005), and writing (U.S. Department of Education, 2002).

Identification of SLD

When the federal regulations for Public Law (PL) 94–142 (1977, the precursor to the Individuals with Disabilities in Education Act) were written, there was no agreed upon diagnostic approach for SLD. The *Illinois Test of Psycholinguistic Ability* (Kirk, McCarthy, & Kirk, 1961) was the most commonly used approach to diagnose SLD during the 1960s and early 1970s, but it quickly fell out of favor due to concerns about the psychometric adequacy of the data (Hammill & Larsen, 1974; Mann, 1971; Ysseldyke & Salvia, 1974). Thus, when SLD became institutionalized in the 1977 regulations for PL 94–142, there was no agreed upon diagnostic criteria and the now-infamous discrepancy model was included in the regulations as a compromise (Gresham et al., 2005).

Research in the 30 years since then has questioned the validity of the discrepancy model on the basis of discriminant validity (i.e., lack of differences between students identified as SLD and struggling readers; Stuebing et al., 2002), consequential validity (i.e., outcomes are not enhanced by diagnosis and services in special education; Algozzine & Ysseldyke, 1983), and social inequity (i.e., disproportionality of and rapidly growing incidence of SLD diagnosis; Ysseldyke, Algozzine, & Epps, 1983). In fact, some have argued against IQ testing as part of SLD identification because of a lack of instructional relevance (Gresham & Witt, 1997; Siegel, 1988) and inability to discriminate readers who require intervention and those who do not (Vellutino, Scanlon, & Lyon, 2000).

More recent data demonstrating the effect of intervention on brain development questions the need for SLD diagnosis and suggests the simplest and most efficient route is to focus on delivering intervention to young students who are struggling to learn. Simos and colleagues (2001) conducted neurological imaging studies and found pre-intervention brain patterns of children identified with SLD that were consistent with an SLD diagnosis (i.e., focusing on the right hemisphere of the brain as they read or no clear pattern). However, after an 8-week intervention, the left cerebral hemisphere showed activity when they read, which was a normalized pattern, and suggested that intervention can be effective in remediating significant learning difficulties.

CAUTION

Research has prompted serious concerns about the validity of the discrepancy model on the basis of discriminant validity (i.e., lack of differences between students identified as SLD and struggling readers), consequential validity (i.e., outcomes are not enhanced by diagnosis and services in special education), and social inequity (i.e., disproportionality of and rapidly growing incidence of SLD diagnosis).

Some have suggested that the discrepancy model is not the best approach to diagnose SLD (Hale, Naglieri, Kaufman, & Kavale, 2004), and others argue that the construct of SLD is fundamentally flawed and will never be adequately conceptualized for identification purposes (Algozzine & Ysseldyke, 1982, Algozzine and Ysseldyke, 1983; Coles, 1998; Ysseldyke & Marston, 2000). Perhaps the only thing that can be stated with any confidence is that, once again, much like in 1977, there are no universally accepted diagnostic criteria for SLD.

POSITIVE FINDINGS OF EARLY RESEARCH

The consistent message from the positive effects of intervention research is that students with reading difficulties can learn at an acceptable rate with quality instruction and that SLD diagnosis can be prevented. Research consistently has demonstrated that instruction and intervention can prevent SLD diagnosis in later years (Lennon & Slesinski, 1999; Torgesen, Rose, Lindamood, Conway, & Garvan, 1999; Torgesen et al., 2001). For example, the seminal study by Foorman, Francis, and Fletcher (1998) compared three instructional strategies among 285 first- and second-grade students who were at risk for reading failure. They found that students who received direct instruction in letter sounds learned reading skills more quickly and had a lower rate of subsequent reading difficulties and SLD than those who were taught sound-symbol relationships by embedding them in connected text or those for whom this instruction was implicitly taught.

Rapid Reference 1.1

Students who received direct instruction in letter sounds:

- Learned reading skills more quickly.
- Had a lower rate of subsequent reading difficulties and SLD.

In addition to preventing SLD diagnosis, effective intervention can lead to positive reading gains even among children with severe reading difficulties and disabilities (Lovett, Borden, Lacerenza, Benson, & Brackstone, 1994; McGuinness, McGuinness, & McGuinness, 1996; Wise, Ring, & Olson, 1999). Some of the components of effective interventions for children with severe reading difficulties include making decisions with formative evaluation (Fuchs & Fuchs, 1986), delivering instruction in small interactive groups (Vaughn, Gersten, & Chard, 2000), and various instructional components such as drill-repetition-practice-feedback, controlling task difficulty, and directed response/questioning

(Swanson, 1999). Moreover, meta-analytic research among students with SLD found large effects for several interventions including mnemonic strategies, explicit instruction, and instruction in comprehension strategies (Kavale & Forness, 1999).

Given the documented poor outcomes associated with special education particularly for children diagnosed with SLD, it seems that preventing learning difficulties is superior to treating them. Moreover, the long history of cultural biases in special education and SLD diagnostic practices (Donovan & Cross, 2002) suggested that alternatives were needed.

RTI DEFINITION AND PURPOSE

RtI is not a new concept (Fuchs & Fuchs, 1998; Velluntino et al., 1996), but it was included in federal legislation only recently, and there seems to be considerable confusion about its implementation (Fuchs, Mock, Morgan, & Young, 2003). The National Association of State Directors of Special Education defined RtI as the practice of providing high-quality instruction, changing instruction based on frequent progress monitoring, and making important educational decisions based on student response to the changed instruction/intervention (Batsche et al., 2005). Others have conceptualized RtI as the systematic use of data to most efficiently allocate resources to enhance outcomes for all students (Burns & VanDerHeyden, 2006). Thus, the commonly described components of RtI are: (a) quality core instruction; (b) universal screening; (c) progress monitoring for students identified with difficulties; (d) increasingly intensive interventions implemented based on student need; and (e) resulting data used to make instructional, resource allocation, placement, and special education identification decisions.

Rapid Reference 1.2

Components of RtI are:

- Quality core instruction
- Universal screening
- Progress monitoring for students identified with difficulties
- Increasingly intensive interventions implemented based on student need
- Resulting data used to make instructional, resource allocation, placement, and special education identification decisions

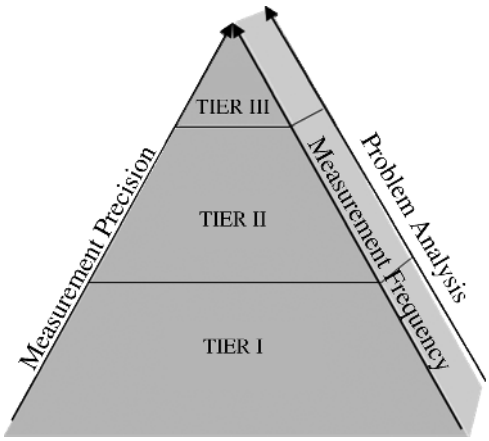

Figure 1.1 Measurement and Problem Analysis Within a Response to Intervention Model.

Source: Based on Burns & Gibbons (2008).

RtI models include multiple tiers of service delivery with most including three tiers. As displayed in Figure 1.1 (Burns & Gibbons, 2008), three things happen as students' needs become more intense and they progress through the tiers. Measurement becomes more frequent and precise, and problem analyses become more detailed and costly. Information about the three tiers on these three issues follows.

DON'T FORGET

It is better to prevent learning difficulties than to treat them, and early intervention shows potential to prevent learning difficulties that otherwise might lead to an SLD diagnosis.

Tier 1

The first tier of any RtI model is quality core instruction. It would be beyond the scope of this book to discuss what constitutes quality reading and math instruction, but assessing the quality of core instruction is a prerequisite to any effective RtI model. Measurement in Tier 1 usually is based on general outcome measures (GOMs), which are essentially general assessments of a student's overall academic performance. For this reason, these assessments often are referred to as the vital signs of learning in that they can be used to reflect in a meaningful way whether children are at risk or not in their instructional programs. Measurement in an RtI system usually relies on general outcome

measures referred to as curriculum-based measurements of reading (CBM-R) and math (CBM-M) because they are sensitive to growth and psychometrically adequate for most decisions (e.g., determining who is at risk, evaluating effects of instruction; National Center on Response to Intervention, 2009). However, Tier 1 assessments may be conducted with less sensitive, but highly reliable, group- or computer-administered tools, such as the Measures of Academic Progress (Northwest Evaluation Association, 2003), Star Math (Renaissance Learning, 1998), and Star Reading (Renaissance Learning, 2003). The goal of assessments within Tier 1 is that they adequately identify a student as proficient in required skills or as needing additional intervention and that they do so in the most efficient way possible (i.e., at the lowest cost to instructional time).

The assessments used in Tier 1 usually are conducted three times each year. Certainly some measures could be conducted more frequently as resources allow and as the data warrant. However, data collected as part of the school's resource allocation system probably should be collected no less than three times each academic year, usually within the first month of school, sometime in January, and again within the last month of the school year.

Because the data collected in Tier 1 are designed to inform a screening or risk decision and are collected somewhat infrequently, only low-level problem analyses are possible. Essentially, the two primary purposes for collecting data within Tier 1 are to (a) identify students who need additional intervention and (b) determine if the problem is specific to the student or the student's classroom. VanDerHeyden and colleagues (VanDerHeyden & Burns, 2005; VanDerHeyden, Witt, & Gilbertson, 2007; VanDerHeyden, Witt, & Naquin, 2003) have consistently demonstrated that class-wide interventions are more efficient and effective when a large portion of the students in one classroom or grade require intervention. In other words, sometimes it is more efficient to take the intervention to the classroom than it is to take students to an intervention. Moreover, quality core instruction is the basis from which all interventions occur. Interventions have a greater likelihood of success if they are highly (and correctly) targeted (Burns, VanDerHeyden, & Boice, 2008), but students learn the skill only if the intervention is contextualized in the broader curriculum. Take reading, for example. A struggling elementary-age reader probably would benefit from additional explicit instruction in sound-symbol relationships, but children cannot be taught how to read simply by learning how to sound out words. The intervention will work only if it is integrated with quality core instruction; without good teaching and curriculum, little else matters.

CAUTION

Class-wide interventions are an efficient and effective way to improve learning where many students are struggling. However, supplemental intervention will work only if it is integrated with quality core instruction. Without good teaching and curriculum, little else matters.

Tier 2

Where Tier 1 instruction is adequate, estimates suggest that up to 20% of students will not be successful despite quality core curriculum and instruction (Burns, Appleton, & Stehouwer, 2005). A Tier 2 intervention is implemented for children identified as struggling learners and for whom a class-wide intervention was either not needed or after it has improved the skills of most students in the classroom. Tier 2 interventions usually are delivered in small groups of 2 to 8 (5 being most common) in elementary school, approximately 8 to 10 for middle school grades, and 10 to 12 or even 15 with high school students. Students are flexibly and fluidly grouped in homogeneous groups based on baseline and progress monitoring data. For example, students who need additional instruction in sound-symbol relationships are in one group, those who need phonemic awareness instruction are in another, and fluency building groups could be in a third. Intervention sessions can be conducted effectively approximately 30 minutes each session 3 to 5 days per week.

Measurement at Tier 2 focuses on more detailed quantification of the learning or performance deficit. Whereas Tier 1 data are used to make a screening decision, Tier 2 data are needed to determine what prerequisite skills are missing and what instructional conditions might accelerate learning (e.g., Does the student demonstrate phonemic awareness? How well does she decode words? How fluent is her reading? and How well does she comprehend what she reads?). These data are used to create homogenous skill groups in order to match the intervention to student need. In addition to being more precise than Tier 1 data, Tier 2 data are collected more frequently. Assessments in Tier 1 occur three times each year, but data are collected in Tier 2 once each week or no less than once every other week. These more frequently collected data are used to monitor progress, to move children between groups, and to judge the effectiveness of the intervention.

Tier 3

On average, 2% to 5% of the student population will require intervention intensity beyond what is provided in Tier 2 (Burns et al., 2005). For those

students, interventions are highly targeted, are developed based on individual student need, and are often delivered in 1-to-1 (1 child to 1 adult) or 2-to-1 formats. Thus, assessment data must go well beyond simply determining how proficient students are in the skill; they must also identify specific skills and skill components that the student knows and does not know. For example, a Tier 1 math assessment would identify a struggling math student and identify a class-wide problem; data collected in Tier 2 would identify automaticity of single-digit multiplication facts as the most appropriate intervention target; but Tier 3 assessments could determine if the student has conceptual knowledge of multiplication (e.g., can use multiplication to find a least common denominator) and which facts the individual student knows and does not know.

DON'T FORGET

Data at Tier 1 are used to make a screening decision. Data at Tier 2 are needed to:

- determine what prerequisite skills are missing and what instructional conditions might accelerate learning,
- monitor intervention progress,
- move children between groups, and
- judge the effectiveness of the intervention.

Data at Tier 3 are needed to:

- build an intervention that will accelerate learning if correctly implemented,
- monitor intervention progress,
- address and adjust integrity and intervention facets to ensure maximal effects, and
- evaluate the intervention effects.

Following the pattern of increased precision and frequency, data are collected in Tier 3 at least once each week to monitor progress. Progress monitoring data collected in Tiers 2 and 3 are often general outcome measures or CBMs (e.g., oral reading fluency and digits correct per minute on a multiskill math probe), but progress should also be monitored in the specific skill being taught (e.g., nonsense-word fluency for a phonics intervention or single-digit multiplication probes). However, even more precise data are used to determine the appropriate intervention and often take into account factors such as the accuracy with which a skill is completed and malleable environmental factors that could contribute to the problem, such as instruction, curriculum, learning environment, and learner characteristics (Hosp, 2008). Some have suggested

testing various interventions to determine which is most successful for an individual student (Daly, Witt, Marens, & Dool, 1997) and using those data to build the intervention (Barnett, Daly, Jones, & Lentz, 2004). In other words, Tier 3 data should help identify the cause of poor academic performance. The purpose of assessment at Tier 3 is to identify an intervention that will accelerate learning when it is delivered before resources are dedicated to deploying that intervention in the classroom.

EARLY IMPLEMENTATION MODELS

Whereas the roots of RtI can be traced to multiple events and literatures, looking back, certain events were seminal for RtI. The University of Minnesota's Institute for Research on Learning Disabilities (IRLD) in the late 1970s greatly influenced the development of what later came to be called RtI. The IRLD conducted groundbreaking research on SLD diagnosis that systematically examined the foundations of SLD diagnosis and service delivery. Those studies caused earthquake effects to the basis and purpose of SLD diagnosis and created impetus for more direct services that would advance student outcomes. Deno and Mirkin's (1977) *Data-Based Program Modification* (DBPM) manual operationalized a problem-solving model for identifying and responding to student learning problems using brief timed measures (CBMs) to inform and evaluate instructional efforts. This important little manual organized and advanced the work of Bloom, Hastings, and Madaus (1971) and those in the precision teaching field, especially Starlin and Starlin (1973), and started the firestorm of research and development on curriculum-based or general outcome measurement. Deno and Mirkin (1977) operationalized a problem-solving process that became the basis for both early and contemporary implementation efforts in what later was referred to as RtI. In the measurement arena, visionary researchers were beginning to raise the idea of contextualized assessment and accurate decision making (Dawes, Faust, & Meehl, 1989) and promoting the idea of treatment utility and consequential validity as a basis for psychological measurement (Messick, 1995).

It is commonly said that we stand on the shoulder of giants, and it is true within RtI implementation as well. Although most implementation initiatives occurred within the past 5 years, a handful of districts and state agencies were engaged in RtI activities decades before RtI was included in federal regulations. Fuchs et al. (2003) identified four major models to which many of the current RtI models can be traced. Brief descriptions of the four models identified by Fuchs et al. are presented next.

The Heartland Area Educational Agency 11 (Heartland) implemented a four-level problem-solving model in 1985 (Ikeda & Gustafson, 2002). Levels I and II involved educational professionals consulting with the child's parents (Level I) and then the school's assistance team (BAT) (Level II). Intervention efforts in Levels I and II were implemented exclusively by school personnel; Heartland staff did not become involved until Level III, at which time teams worked with school personnel in an extended problem-solving process. Finally, students for whom intervention efforts in Level III were not successful were considered for special education eligibility (Level IV). Heartland recently transitioned to a three-tier model (Tilly, 2003) and remains one of the best-known and most well-respected RtI implementation sites in the country.

Minneapolis Public Schools (MPS) also implemented a problem-solving model (PSM) in 1993 that merged special and general education personnel (Marston, Muyskens, Lau, & Canter, 2003). The PSM closely monitored student progress, accommodated students in general education, and provided a non-biased method of identifying children in need of special education (MPS, 2001). There are three stages in the PSM that progress from teacher classroom interventions based on universal screenings (Stage 1), to refined interventions and progress-monitoring strategies developed by a problem-solving team (Stage 2), and consideration of special education referral in Stage 3 (Marston et al., 2003). Although school districts across the country only recently have begun using RtI data for eligibility decisions, MPS did so through a state waiver over 15 years ago and was among the first to do so.

Pennsylvania's Instructional Support Team (IST) model was phased into all elementary schools of the state's 501 school districts over a 5-year period that began in 1990 (Kovaleski, Tucker, & Duffy, 1995). The model was implemented in an attempt to bridge special and general education programs by shifting the focus of special education from categorical services to effective instruction in general education (Kovaleski et al., 1996). The primary component of the IST model was the instructional support teacher who was specially trained and who worked exclusively with classroom general education teachers to assist with struggling learners (Kovaleski et al., 1995). However, the instructional support teacher did not deliver direct support to any students beyond modeling instructional approaches for the classroom teacher and occasional short-term interventions. The instructional support teacher could provide support for 50 school days, at which time the IST met to discuss student progress and decide if the student would be referred for a multidisciplinary evaluation for special education eligibility.

There were no formal phases or stages within IST, but three basic steps were followed:

1. An initial consultation took place between the classroom teacher and a consulting member of the IST.
2. Teacher concerns were behaviorally defined and the IST was convened.
3. The IST developed interventions that were collaboratively implemented by the classroom teacher and the support teacher (Pawlowski, 2001). IST was described as "the best-known statewide pre-referral intervention program in the nation" (Fuchs et al., 2003, p. 162). In addition to Pennsylvania, it was implemented on much smaller scales in Connecticut, Michigan, New York, and Virginia.

Ohio's statewide Intervention-Based Assessment (IBA) model emphasized functional and direct assessments of academic difficulties to identify and evaluate interventions for student learning and behavioral difficulties (Barnett et al., 1999). The multidisciplinary team (MDT) consisted of educational professionals and the child's parents and relied heavily on conjoint behavioral consultation (Telzrow et al., 2000). There were no specific phases in IBA and no mandated timelines. However, the MDT could conduct a special education eligibility evaluation at any time in the process if instructional methods necessary for success represented specially designed instruction, the child's characteristics matched the federal definition of one or more special education disabilities, and it was determined that the condition would have had an adverse effect on the child's education without special education and related services (McNamara & Hollinger, 2003). Much like IST, IBA was essentially a prereferral intervention process.

CURRENT PRACTICE AND RESEARCH

RtI has moved from isolated islands to a widespread network with a continuum of implementation progress. As stated earlier, there are RtI implementation sites in all 50 states, but Jimerson, Burns, and VanDerHeyden (2007) identified six RtI models, in addition to the four aforementioned ones, that were a second-wave of implementation leaders. These leaders include the St. Croix River Education District in Minnesota; the Illinois Flexible Service Delivery model; the System to Enhance Educational Performance (STEEP) operating in districts in several states; Michigan's Integrated Behavior and Learning Support Initiative; Idaho's Results-Based Model; and Florida's Problem-Solving statewide model. Clearly the practice is widespread and growing.

Recent research supports the recent growth in RtI initiatives. Specifically, implementing an RtI model resulted in more children demonstrating proficient skills on state accountability tests (Heartland, 2004; Sornson, Frost, & Burns, 2005; VanDerHeyden & Burns, 2005), improved reading skills among children identified as at-risk for reading failure (Marston et al., 2003; Tilly, 2003), more accurate and equitable identification of students in need of special education (VanDerHeyden et al., 2003; VanDerHeyden & Witt, 2005), and fewer children being placed into special education (Burns et al., 2005; Sornson et al., 2005; VanDerHeyden et al., 2007). Approximately 5.7% of school-age children were identified with a SLD (United States Department of Education, 2002), but fewer than 2% of the student population in various studies and program evaluations of RtI were identified with SLD (Burns et al., 2005). Perhaps one of the most comprehensive and experimentally rigorous evaluations of a multitiered intervention model found that students increased reading achievement and the percentage of new placements in special education fell from 15% to 8% (O'Connor, Harty, & Fulmer, 2005). Moreover, children with identified disabilities in an RtI model received more services and additional specialized instruction as compared to more traditional approaches (Ikeda & Gustafson, 2002; Reschly & Starkweather, 1997), and special education services began at earlier grades (Reschly & Starkweather, 1997).

Schools and school districts have only recently begun moving RtI principles and procedures to middle and high schools, and research thus far has been scant. Vaughn et al. (in press) implemented a large-scale RtI initiative with seven middle schools and examined the effectiveness of a Tier 2 intervention. Those students who received the intervention outperformed the control group on several reading measures including word attack, comprehension, and phonemic decoding, but the effects were small. Implementation at the high school level is probably even rarer than middle school, but Burns (2008) describes a model that currently is used in practice, and previous efforts resulted in positive outcomes with reading scores of participating ninth-grade students growing at a rate three times that of typical students and more than five times greater than their own growth in eighth grade (Windram, Scierka, & Silberglitt, 2007). Although additional research is needed, there is an evidence base from which to build RtI implementation efforts.

SUMMARY

The movement toward RtI began in 1977 with the publication of the *Data-Based Program Modification* manual (Deno & Mirkin, 1977), and subsequent research has

suggested positive effects. One of the key components of effective intervention is the use of formative evaluation (Fuchs & Fuchs, 1986) and assessment data to determine which interventions would have the highest likelihood of success (Burns et al., 2008). Thus, RtI again has emphasized the importance of assessment data within the instructional process (Gresham, 2002) and renewed the debate about what constitutes instructionally relevant assessment data (Batsche, Kavale, & Kovaleski, 2006; Gresham et al., 2004; Hale et al., 2004).

In our opinion, again returning to the wisdom of Reynolds (1975): "[In] today's context the measurement technologies ought to become integral parts of instruction designed to make a difference in the lives of children and not just a prediction about their lives" (p. 15). The vision for assessment written by Reynolds and Stiggins (2005) and others has not yet reached fruition for a number of reasons. However, RtI presents perhaps the best opportunity for educational assessment to reach its potential since 1905 when Alfred Binet translated his test into English. Positive outcomes have been found in early implementation, but only when specific implementation procedures, assessment practices, and decision rules were used.

TEST-YOURSELF QUESTIONS

1. **Why does a system like RtI have diagnostic implications for students?**
2. **Historically, SLD identification practices have been criticized as having poor consequential validity. What is the basis for this criticism?**
3. **Define RtI. What are the key components needed for RtI implementation?**
4. **What decisions are made based on assessment data collected at Tier 1?**
5. **Does it make sense to implement Tier 2 intervention in the face of inadequate Tier 1 or core instructional practices?**
6. **What percentage of students may be expected to need Tier 2 intervention? What percentage of students may be expected to need Tier 3 intervention?**
7. **What research evidence suggests that RtI can be used to reach decisions that have consequential validity for students and systems?**

Answers:

1. Because RtI use creates a data set that can be used to determine eligibility for services under the category of SLD.
2. There is no evidence that making the diagnosis leads to treatment that enhances outcomes. Traditionally there has been poor classification agreement and little basis for discriminating between those with poor academic skills and those with SLD.
3. RtI is a system of decision making to allocate instructional resources to enhance learning outcomes for all students. Components include quality core instruction, progress monitoring for students below criterion, increasingly intensive interventions implemented based on student need, and resulting data used to allocate resources and make special education eligibility decisions.

(*continued*)

4. Identification of system and individual learning problems (i.e., screening), evaluation of adequacy of core instruction.
5. No.
6. Up to 20% may need Tier 2 intervention; 5% to 10% may need Tier 3 intervention.
7. With properly implemented RtI, student learning improves and diagnosis even may be prevented.