



Response to Intervention in Literacy: Problems and Possibilities

Author(s): Peter H. Johnston

Source: *The Elementary School Journal*, Vol. 111, No. 4 (June 2011), pp. 511-534

Published by: [The University of Chicago Press](http://www.press.uchicago.edu)

Stable URL: <http://www.jstor.org/stable/10.1086/659030>

Accessed: 19/07/2011 12:46

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=ucpress>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *The Elementary School Journal*.

RESPONSE TO INTERVENTION IN LITERACY

Problems and Possibilities

ABSTRACT

Evidence suggests that it is possible to substantially reduce the number of children classified as learning disabled in literacy. The 2004 reauthorization of the Individuals with Disabilities Education Act opens opportunities to do so through Response to Intervention (RTI). This article argues that institutional histories, structures, and professional belief systems are leading to patterns of implementation that emphasize the need to identify individuals with disabilities rather than reducing the need for such identification. Rapidly becoming the default, these approaches to RTI do not fully support the aspiration of the law. This article documents the beliefs, assumptions, institutional structures, and divisions of expertise that have led to and sustain this focus on identification. It also explores a contrasting frame emphasizing instruction and prevention and the associated views of literacy teaching and learning. Finally, the law's requirement of "research-based" literacy instruction is examined and a way forward proposed.

Peter H. Johnston

UNIVERSITY AT ALBANY,
STATE UNIVERSITY OF
NEW YORK

HOW should we understand a child's limited progress in acquiring literacy? We have laws and institutional structures reflecting a belief that some children have intrinsic factors impeding their acquisition of literacy and that these children require a separate system of "special" education. Researchers point to genetic histories (Harold et al., 2006), brain scans (Backes et al., 2002; Hoeft et al., 2007), and extensive cognitive assessments (Catts, Hogan, & Fey, 2003; Vellutino et al., 1996) to support the existence of these individual predisposi-

tions. There is little doubt that when children come to school there are differences in their experiences with language and literacy (Heath, 1983; Pancsofar & Vernon-Feagans, 2006). There are also differences in their facility with various language-processing competencies that make it harder for some to acquire literacy (Vellutino et al., 1996). But we should not lose sight of some important facts. First, these differences are not bimodal distributions with a clearly separable group of students with specific learning disabilities (SLD). Just as with measures of height or weight, or intelligence, there is no clear line of demarcation between SLD students and others. Second, with appropriate instruction, even students who are most limited in these competencies can be taught to read (Scanlon & Vellutino, 1997). Third, as these children learn to read, the measures of cognitive processing on which they previously appeared deficient improve as well and no longer look as disabling (Vellutino et al., 1996). Fourth, although 1%–2% of students (Al Otaiba & Torgesen, 2007; Clay, 1990; Vellutino et al., 1996) still have difficulty even after these generally successful interventions, it is possible (though not yet common) to bring three-quarters of even these students up to par with their peers by fine-tuning their instruction (Phillips & Smith, 1997).

Focusing on particular studies, and not explaining what is meant by “taught to read,” oversimplifies a more complex problem; nonetheless, we have much of the knowledge necessary to prevent most children from becoming disabled in literacy. This is the aspiration of Response to Intervention (RTI), a provision in the 2004 revision of the federal Individuals with Disabilities Education Act (IDEA) legislation (Assistance to States for the Education of Children with Disabilities, 2006).

In this article, I will argue that current patterns of RTI implementation in literacy do not fully support the aspiration of the law. These implementation patterns, which are rapidly becoming the default approach to RTI, emphasize the need to identify individuals with disabilities rather than emphasizing reducing the need for such identification. Beginning with a little historical context, I describe what it means to focus on identification, emphasizing measurement at the expense of instruction. I document the associated beliefs, assumptions, institutional structures, and divisions of expertise that have led to and sustain this focus; I also explore a contrasting frame emphasizing instruction and prevention. Then, in light of this discussion, I explore RTI assessment requirements and practices and the associated views of literacy teaching and learning. Finally, the law’s requirement of “research based” literacy instruction is examined and a way forward proposed.

RTI in the History of SLD

Explaining why some children encounter difficulties in learning, and finding what to do about it, has consumed considerable time and resources, particularly when the children are middle class and intelligent. In 1963 a group of parents with influence organized a conference to find suitable explanations and to seek resources to remedy these children’s problems. There were two outcomes of their efforts: (a) the designation “specific learning disability” (SLD), a more acceptable alternative to the tendency to equate low achievement, particularly in reading, with low intelligence, and (b) funding to support special instruction for children with adequate intelligence but difficulty with academic tasks, such as learning to read. Passage of the Education for All Handicapped Children Act (EAHCA) in 1975 (reauthorized as IDEA in 1990)

secured these outcomes. “Specific learning disability” acquired its current definition as:

(i) *General*. Specific learning disability means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations, including conditions such as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia.

(ii) *Disorders not included*. Specific learning disability does not include learning problems that are primarily the result of visual, hearing, or motor disabilities, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage. (Assistance to States for the Education of Children with Disabilities, 2006, p. 46757)

IQ Testing and SLDs

In practice, there was no way to tell whether a person had an SLD except by exclusion. If other factors failed to explain a lack of success, the cause was probably an SLD. Consequently, if IQ test scores showed a child to be intellectually capable yet not demonstrating a comparable level of reading ability, the child was considered learning disabled in reading unless he or she was poor or a minority, or had some other apparently explanatory impediment such as blindness or deafness. Indeed, presumably because of these exclusionary factors, studies of learning disabled (LD) students between 1963 and 1973 reported largely middle-class students, with only 1.5% minorities (Sleeter, 1986).

IQ testing was the province of school psychologists, and this expertise centralized them in the process of identifying SLD. However, in the years leading up to the 2004 IDEA revision, the IQ-achievement discrepancy came into question (Aaron, 1997; Fletcher, 1992; Stanovich, 1991; Vellutino, Fletcher, Snowling, & Scanlon, 2004). First, IQ tests turned out to be questionable indicators of general ability, particularly for cultural and linguistic minorities and low socioeconomic status (SES) children (Beiser & Gotowiec, 2000). Second, for children with lower IQs, limited literacy development could be seen as “normal” and thus acceptable. Third, for many children it took a couple of years before the discrepancy between normative achievement and IQ became sufficiently substantial, by which time a considerable history of failure and confusion could accumulate. Fourth, this testing provided no instructionally useful information (Vaughn, Levy, Coleman, & Bos, 2002). Fifth, the discrepancy does not predict how well students respond to an intervention (Vellutino, Scanlon, & Lyon, 2000). A paper by influential researchers argued persuasively for eliminating the IQ discrepancy definition (Lyon et al., 2001), and the initial draft of the law allowed states to prohibit its use, citing numerous studies (Assistance to States for the Education of Children with Disabilities, 2006, p. 46543). Nonetheless, the crafters of the IDEA revision chose not to eliminate the discrepancy definition. Instead, they introduced an alternative approach to the problem, now referred to as Response to Intervention, or RTI.

Costs and Benefits of SLD Classification

Being classified as having an SLD brought extra resources through special education such as a reduced teacher-pupil ratio. However, there were also drawbacks to the classification. First, by definition, the problem was viewed as fundamental (“a disorder in one or more of the basic psychological processes”; Assistance to States for the Education of Children with Disabilities, 2006, p. 46757), so instruction was provided not to eliminate the disability but the symptoms. Second, special education teachers, trained to work with a broad range of atypical students, in many cases received less training in literacy instruction than did regular classroom teachers (Jakupcak, 1975). Consequently, the label brought more expensive, though not necessarily more expert and not particularly effective, resources (Kavale, 1990; Vaughn et al., 2002). Relatively few children exit special education (Carlson & Parshall, 1996), and for those who do there is no clear link to special education practices.

Growing Numbers of SLD Students

With recognition of SLD and funding linked to it, the category rapidly acquired members. In the 1990s, preceding the initiation of the law, there was a 34% increase in the number of schoolchildren classified as SLD (Advocacy Institute, 2002). Although this rate has since leveled off (Westat, 2009), in some districts over 15% of students are classified as SLD (U.S. Department of Education, 2002). Increases in accountability testing also fueled this increase in the number of students labeled as SLD. Because SLD was assumed to be permanent, schools could not be held responsible for the normal development of these students, so they were not included in large-scale accountability testing. This, along with the special education money that accompanied classification, provided an incentive for classification (Allington & McGill-Franzen, 1992). This made it possible to remove many low-achieving students from the accountability roll; disproportionate numbers of these students were cultural minorities and in poverty (Arnold & Lassmann, 2003). Indeed, whereas minorities originally made up only 1.5% of the special education population, currently over 20% of the national special education population is African American, compared to less than 14% of the general population (Elementary and Middle Schools Technical Assistance Center, 2009).

Institutional Investments and RTI

It is fair to say that RTI was a rational response to at least the four factors described above: (a) the demonstrable problems with IQ-achievement discrepancy assessments, (b) the burgeoning number of students classified as having an SLD (with associated costs), (c) the documented effectiveness of early intervention efforts for reducing that number (and associated costs), and (d) the overrepresentation of minority students in special education. However, RTI is also framed to extend the reach of the Reading First component of the No Child Left Behind (NCLB) Act, particularly the reduction of special education and the enforcement of “scientifically based” instruction.

With respect to special education, RTI represents a major shift in institutional priorities that potentially destabilizes existing institutional investments, creating awkward tensions. One set of tensions has arisen from the potential shift in funding

built into the law. Because improving the quality of instruction and adding high-quality interventions require resources, the crafters of the 2004 reauthorization of IDEA allowed up to 15% of the special education budget to be used for this purpose. On the one hand, then, the logic for RTI is based in regular education: before you classify a child as disabled, ensure the problem is not an instructional one. On the other hand, the RTI funding stream, intended to reduce the need for special education, is in the special education budget. If states and schools productively and aggressively take up the RTI option, it will reduce the funding and the need for special education teachers.

A second set of tensions has arisen in the school psychology community because of the reduced need for IQ testing and increased need for expertise in literacy teaching and learning. As the “costs and benefits” section (under Executive Order 12866) of the final regulations in the *Federal Register* (Assistance to States for the Education of Children with Disabilities, 2006) recognizes, RTI eliminates part of school psychologists’ historical practice—IQ testing to identify SLDs (see particularly pp. 46746, 46750, and 46751). School psychologists report that up to 75% of their time has been spent classifying children into special education, principally through SLDs (Reschly & Wilson, 1995), and that 79% of referrals include reading disability (Nelson & Machek, 2007). The vast majority of students classified as SLD are referred for problems in literacy. In Florida, the figure is 95% (Batsche, Curtis, Dorman, Castillo, & Porter, 2007). The majority of school psychologists, however, report little knowledge of reading assessment and interventions (Nelson & Machek, 2007), let alone literacy more broadly defined. With some exceptions, there is no space in their expansive training programs to develop detailed expertise in literacy teaching and learning even to the level expected of classroom teachers. The National Association of School Psychologists (NASP, 2000), which accredits programs, does not require program applicants to have teaching certification or experience; NASP lists 11 broad essential-knowledge areas, only one of which could accommodate literacy teaching and learning, “effective instruction and development of cognitive/academic skills” (2000, p. 15).

Perhaps because RTI is in the IDEA law, it is popularly seen as centrally about children with disabilities and thus the province of school psychologists and special educators. Consequently, the initial flood of books and articles on the subject have virtually all been written by authors from these two fields. In the first handbook on the subject (Jimerson, Burns, & VanDerHeyden, 2007b), 53 of the 86 authors are from school psychology and 18 are from special education. (The rest are from psychology or administration, or are not identifiable.) A search of EBSCO for RTI or Response to Intervention in the title produces a ratio of school psychologists and special educators to regular education authors of between 15 and 20 to 1 (with some mixed authorship). The federally funded RTI Action Network Web site lists 42 contributors, 21 from special education, 13 from school psychology, and 8 from different areas such as speech and language pathology, neuropsychology, and administration. These are the dominant voices in the conversations about RTI and as such stand to influence policy and practice. Though there are differences among these writers, on the whole both professional groups appear to share a belief in the permanence of SLDs, a behaviorist view of literacy teaching and learning that emphasizes the speed and accuracy of reading words, a belief that it is reasonable to assess literacy development as if it were linear, a central concern for psychometrics, and a positivistic

view of science. I will argue that this constellation of commitments and beliefs, along with the prominence of the professional voices, is highly problematic.

RTI as Identification or Prevention?

IDEA introduces RTI by requiring that states “must permit the use of a process based on the child’s response to scientific, research-based intervention” (Assistance to States for the Education of Children with Disabilities, 2006, p. 46786) as part of addressing the problem of SLD. The requirement occurs as both an alternative assessment to the IQ discrepancy identification strategy (§300.307), and as part of ensuring that a child has received “appropriate instruction” (§300.309). For example, §300.307 of the law is about “identifying” SLD, but in the process, it must be determined “that underachievement . . . is not due to lack of appropriate instruction in reading or math.” This requires that the responsible parties examine (a) “data that demonstrate that prior to, or as a part of, the referral process, the child was provided appropriate instruction in regular education settings, delivered by qualified personnel” and (b) “data-based documentation of repeated assessments of achievement at reasonable intervals, reflecting formal assessment of student progress during instruction” (Assistance to States for the Education of Children with Disabilities, 2006, p. 46787).

In other words, RTI can be taken up primarily as identification, essentially a measurement problem, or primarily as prevention, an instructional problem. Although identifying an SLD and preventing one (i.e., identifying, instead, how to improve instruction) might be viewed as different parts of the same process, they are different activities requiring different tools and strategies (Fuchs, Stecker, & Fuchs, 2008) and different discourses and relationships. Although it is not a simple dichotomy, these differences along with limits on time and other resources mean that emphasizing one means limiting the other.

RTI as Identification: A Measurement Problem

If the emphasis in RTI is on identifying who does and does not have SLD, it becomes a psychometric or measurement problem—an effort to find a standardized procedure to replace the IQ–achievement test discrepancy procedure. In this frame, standardization is important. A common example of such an approach to RTI, referred to as a “standard protocol” approach (Gresham, 2007), would be to use a “scientific, research-based” phonics package (often scripted), checking to make sure all teachers use it in the standard way (“with fidelity”). If in a standard period of time the child has not improved sufficiently, either the intervention is applied with increased “intensity” or “dosage” (Deshler, 2008)—more time, smaller teacher-pupil ratio—or the child is “identified” as SLD and referred to special education. These steps in instructional escalation prior to classification have become referred to as tiers, with tier 1 being classroom instruction and successive degrees of intensity being tiers 2 and 3, with 3 tiers being the commonly suggested model (Batsche et al., 2005). Federal regulations make no mention of tiers, much less three of them.

The most common associated assessment is frequent measurement, often twice a week (Ardoin, 2006), of the number of words a student reads correctly from a standard text in a minute (CWPM). This has become “standard practice” in school

psychology programs (Olson, Daly, Anderson, Turner, & LeClair, 2007) and common in state implementations. Short interventions are preferred in this measurement-oriented approach because “waiting 10 weeks to learn that an intervention is ineffective is troublesome” (Ardoin, 2006, p. 713). Nonetheless, Ardoin (2006) cautions against short 5-week interventions (suggested by Shinn, 2002) because of “relatively large standard errors of the estimates” and the consequent need for at least 20 (rather than 10) data points for accurate measurement (p. 713). In other words, in this frame measurement rather than instructional concerns govern length of the intervention.

The standard protocol approach is often distinguished from problem-solving models that emphasize efforts to “isolate target skills/subskill deficits and shape targeted interventions” (Jimerson, Burns, & VanDerHeyden, 2007a, p. 4). However, in practice problem-solving approaches commonly have standardized components (Burns & Coolong-Chaffin, 2006), and both virtually always depend on curriculum-based measurement (CBM) for assessment—most commonly CWPM. For example, the state RTI program in Idaho uses CBM and is characterized as invoking standard protocol for the first two levels of intervention, then problem solving (Callender, 2007). Similarly, the statewide initiative in Florida centralizes CBM (Batsche et al., 2007). None of this is required by the law. The regulations make no mention of how frequently progress should be monitored (perhaps three times per year rather than twice a week is more sensible) or by what means. Though some argue that CWPM takes little time and therefore is useful even if it offers no instructionally relevant information, the instructionally relevant information already being collected in many classrooms can accomplish the monitoring function without adding further assessment that does not contribute to instruction (e.g., McGill-Franzen, Payne, & Dennis, 2010). The law does not require separate assessments for instruction and monitoring.

Problematic Assumptions

The standard protocol approach (and related forms of problem solving) rests on several assumptions. First, it assumes that a standardized intervention will transfer effectively to a new setting regardless of age, teacher experience/expertise, context, instructional history, et cetera. Second, it assumes that a practice shown to be effective on average with a group of children will be effective with each child experiencing difficulty in the new context. If the child’s reading improves, it is assumed that the instructional package worked; if not, the child is framed as the problem—a treatment resister (Torgesen, 2000) or a chronic nonresponder (Fuchs et al., 2008). The instructional package remains “scientific and research based.” Third, as is evident in the language just described, it frames the problem as a fixed trait of the child. Fourth, it assumes that the materials, rather than a responsive teacher, are the effective component. Fifth, common applications assume that increasing the “dosage” will increase the effect. Sixth, common applications assume a narrow, behaviorist view of literacy, focusing centrally on accurate and rapid identification of words rather than on a broader definition of literacy. In part, this is because the more complex the view of literacy, the more standardization of instruction becomes transparently problematic and the more reliable assessment becomes difficult.

There are good reasons to be cautious about these assumptions. For example, a study by Wanzek and Vaughn (2008) offers evidence that, at least with low-performing first graders, we can assume neither that increasing the dosage will increase the effectiveness nor that instruction that is effective on average is effective for the individual. These researchers used a standard protocol intervention, implemented with high fidelity, and doubled its intensity for one group of students. They found a modest effect for their intervention but no academic benefit for the increased intensity, partly because “tutors reported difficulties throughout the 13-week intervention with student fatigue, group management, and increased problem behavior during the second 30-minute session” (Wanzek & Vaughn, 2008, p. 139). Upon the limited success of their intervention for some students, they concluded that “the likelihood that these students were truly ‘nonresponders’ is quite high. In our view, these students were the very-difficult-to-teach youngsters that a response to intervention model is seeking to identify and consider for special education” (Wanzek & Vaughn, 2008, p. 138). They argue that “students whose response to intervention has been relatively low are likely to require very intensive and ongoing interventions over time, and their response to these interventions is likely to be slow, . . . [these students] may need different instruction than other at risk readers” (Wanzek & Vaughn, 2008, p. 138). They do not say what this different instruction might be like.

Although their intervention improved children’s performance in selected areas on average, Wanzek and Vaughn (2008) took the laudable step of identifying the proportion of students whose competence actually deteriorated by at least half a standard deviation during their interventions. In the single intervention, 14% deteriorated in word identification, and 24% in word attack—the primary targets of the intervention. With the “double dose” intervention, the proportion deteriorating was 7% on word identification, 35% on word attack, and in addition 14% did worse on comprehension. Their analysis shows the questionable practice of taking an instructional approach that is effective on average and assuming that if implemented “with fidelity” on other students it should be effective for each of them, particularly for those experiencing the most difficulty.

Significance of Expertise and Interaction

The measurement frame represents a serious underestimation of the significance of human interaction and expertise in literacy teaching and learning and a problematic view of social science. Indeed, writers from this perspective sometimes classify as standard protocol programs that attend closely to the instructional interaction between teacher and student—the part requiring expertise—because they have a structural component. For example, the most demonstrably successful RTI-related reading interventions so far are Reading Recovery (RR, <http://ies.ed.gov/ncee/wwc/>) and the Interactive Strategies Approach (ISA; Scanlon & Vellutino, 1996; Scanlon, Vellutino, Small, Fanuele, & Sweeney, 2005), both of which have been referred to as standard protocol programs and described as “scripted” because they have consistent lesson segments (Gresham, 2007). Neither program is scripted, and neither would work if it were (Clay, 2005; Vellutino & Scanlon, 2002). For example, suppose during instruction a child reads a word incorrectly. A scripted program would prescribe the teacher’s response. By contrast, in RR and ISA the teacher’s response will depend on, among other things, the text difficulty, the instructional opportunity

offered by the word, the context of the error, and the child's current processing strategies. This means that monitoring for "fidelity" would need to be quite different in scripted and responsive instructional programs.

Treatment fidelity in a scripted program would mean checking to make sure that the teacher stayed on script. However, taking this approach in a responsive teaching program would miss, and in practice, risk discouraging the central component of adaptive teacher expertise. In a responsive teaching program such as RR or ISA, there are certainly programmatic structures, such as a certain amount of time on rereading, writing, reading new material, building word knowledge, and so forth, that could be monitored, but beyond that fidelity would mean adhering to program principles, including responsiveness to children's learning and behavior. Monitoring for fidelity would not only be different, it would require different expertise on the part of the monitor.

In a measurement frame, the valued expertise is not the teacher's ability to adapt instruction but the design and selection of tests and packaged programs. Expertise is minimized by the selection of instructional packages, particularly scripted ones, and testing instruments that can be used by people with limited expertise. Indeed, writing from this perspective, Kovaleski (2007) argued, against the evidence, that tutoring programs "such as Reading Recovery, have produced only moderate gains," but that "impressive results" have been obtained with "commercially available packages" that can be used by "trained instructional assistants" (p. 83). Similarly, in a video conference offered as a resource by the federally funded National Center on Response to Intervention and developed by the National Center for Learning Disabilities (2008), the experts, three school psychologists, described the need for tests that are "simple and quick to use" so you can "use almost anyone" to administer them. They referred to the teams administering these tests as "swat teams," and wondered, "How are we going to provide the technology and data management systems to the teachers and the principals so they can get real live on time data?" The assumption is either that teachers do not, or cannot, keep notes on regular brief conferences with children about their reading and writing, take records of children's reading and writing processes, or collaborate with children to keep systematic files of their work—assessment they could not effectively teach without. Alternatively, the assumption is that the records teachers keep are not adequate for documenting development or informing instruction. But, if teachers are not able to assess in ways that offer instructionally useful information, it is likely that instruction is not tailored to individual children, and there is a need for interventions that increase teacher expertise.

RTI as Prevention: An Instructional Problem

It is possible to view RTI as centrally about ensuring "appropriate instruction" by "qualified personnel" (§300.309)—optimizing instruction to prevent the need to classify children as SLD. This part of the law requires "data that demonstrate that prior to, or as a part of, the referral process, the child was provided appropriate instruction in regular education settings, delivered by qualified personnel." If the emphasis is put on instruction, then evidence that the child is not learning adequately is primarily evidence that instruction is not yet appropriate and needs to be further optimized. It is evidence of a need for greater instructional expertise and perhaps a reduced teacher-student ratio. The fact that instruction was the same as that received

by a group of students in a research study would not be adequate evidence of appropriate instruction for this student. Assessment practices would capitalize on and extend teachers' expertise, particularly their ability to optimize instruction for individual students.

Instructional Expertise

Teacher expertise is the most important factor in improving children's learning (Darling-Hammond & McLaughlin, 1999), and children experiencing the most difficulty should have the most expert teachers (Clay, 1985). In principle, this seems uncontroversial. In practice, however, this is commonly not the case. In the current RTI measurement frame, children can have their standard protocol intervention program delivered by a trained instructional assistant (Kovaleski, 2007) before being sent to special education, where literacy instruction is less than optimal (Vaughn et al., 2002). The key intermediary has been the school psychologist (Reschly & Wilson, 1995), whose expertise does not center on literacy (Nelson & Machek, 2007).

In an instructional frame, expertise in literacy teaching is central, and research on intervention efforts bears this out. For example, in Phillips and Smith's (1997) one-to-one efforts to rescue the lowest 1.5% of literacy learners, the handful of children not recovered were concentrated with particular teachers. Similarly, in Scanlon and her colleagues' work, in some schools none of the intervention students lost ground after the end of the first-grade intervention; in others over 50% did (Scanlon, 2011). The source of this problem is currently under investigation, but evidence suggests that teacher expertise and the institutional structures that support the development of that expertise are central.

Improving teacher expertise has powerful effects. For example, in one successful intervention, Vellutino and colleagues trained special intervention teachers to work with small groups of at-risk kindergarten students, substantially reducing the number of students requiring support in first grade (Vellutino, Scanlon, Small, & Fanuele, 2006). In their next intervention, they taught the existing kindergarten teachers the skills and knowledge to guide their instructional planning and decision making, and to do the small-group work. The effect was the same (Scanlon, Gelzheiser, Vellutino, Schatschneider, & Sweeney, 2008). Others have found similar effects of professional development (McGill-Franzen, Allington, Yokoi, & Brooks, 1999). These are not standard protocol interventions, and they raise the possibility of RTI models that might have continuous improvement for each instructional level built in. Most current models of RTI do not address this in part because their emphasis is on standard protocol intervention packages rather than on teacher expertise, and in part because of the logic of research-based instruction, which commonly renders the program valid regardless of the lack of effect on an individual.

Constant improvement of classroom instruction requires ongoing professional development and probably a well-trained coach (Dorn & Schubert, 2008), a position that requires skill beyond, and including, that of the most skilled classroom teacher. As both the Comprehensive Intervention Model (Dorn & Schubert, 2008) and Reading Recovery programs argue, a productive coach needs to teach children part of the time to maintain skill and respect. In these programs, as with Phillips and Smith's (1997) work, teacher and coach not only discuss data but also take turns doing the teaching in order to explore alternatives. Coaching in this sense is infrequently men-

tioned in special education and school psychology models of RTI. Though interest may be increasing (e.g., Roehrig, Duggar, Moats, Glover, & Mincey, 2008), the 32-chapter *Handbook of Response to Intervention* (Jimerson et al., 2007b) has two references to coaching, one parenthetical and the other a single paragraph. The preferred model is “consultation,” where a teacher seeks the advice of a school psychologist (Knotek, 2007). To effectively coach, however, one must have more knowledge about literacy, teaching, and learning than the teacher and know how to coach.

Assessment Expertise and Practices

In an instructional frame, assessment that informs instruction is key, and teacher assessment expertise is central. A core dimension of teacher expertise is the teacher’s ability to notice and respond to what children can and cannot do. A teacher who does not notice that a child is not yet able to analyze their speech into phonemes will not know to teach that skill. A teacher who does not recognize that a child already has the skill may go ahead and waste time teaching it anyway, particularly if using a packaged program. Similarly, a teacher who notices that a child makes many reading errors and attributes it to the child’s inability rather than to the text’s difficulty will keep the child in an unproductive learning situation.

Instructionally, this is the assessment that matters—what teachers notice and respond to during the literate activities of the classroom. What they notice and what they think it means influences the instruction they provide, unless of course their interactions are scripted and do not take account of children’s current learning and behavior at all. Teachers whose descriptions of their students’ literacy development are brief and lack detail are more likely to refer children for LD evaluation than are teachers whose descriptions are more detailed (Broikou, 1992). Teachers favoring whole class instruction are more likely to refer children for LD evaluations than are teachers who favor a combination of grouping structures (Drame, 2002). Teachers who use small-group instruction in their classes are generally more effective than those who do not (Scanlon et al., 2008; Taylor, Pearson, Clark, & Walpole, 2000). In other words, teachers who do not understand the children they are teaching will be more likely to decide that a child is learning disabled. Their understanding of children’s learning has a lot to do with what they know about children’s literacy acquisition and how to organize to make literate behaviors noticeable and respond to them productively, which is exactly the point of RTI.

When instruction to prevent SLD is first priority, it will not be enough to monitor how much a student is learning. Teachers will need to have instructionally relevant information about the student’s learning, acquired and examined in a way that makes it likely to be used. Collecting and attending to such information in a shared learning environment not only accomplishes change in student performance, it also accomplishes teacher development (Gilmore, 2002). Indeed, research suggests that teacher learning communities focused on data analysis are “the most powerful single approach to improving student achievement” (William, 2006, p. 1). To work for optimal instruction for those students whose learning is most challenging to us, however, we will also need specific information about the nature of the interactions between teacher and student—data on the timing, wording, examples, focus, and control in teaching (Phillips & Smith, 1997). In this frame, the process of assessment

must constantly make personnel, individually and collectively, more instructionally expert.

Assessment in RTI

IDEA requires that a committee evaluating a student for a possible SLD must consider “data-based documentation of repeated assessments of achievement at reasonable intervals, reflecting formal assessment of student progress during instruction” (Assistance to States for the Education of Children with Disabilities, 2006, p. 46787). There are many ways to interpret this directive, particularly regarding what constitutes a reasonable interval, formal assessment, and data-based documentation of achievement. The law also requires the committee to consider “data that demonstrate that prior to, or as a part of, the referral process, the child was provided appropriate instruction in regular education settings” (Assistance to States for the Education of Children with Disabilities, 2006, p. 46787). In other words, monitoring needs to determine whether a child is responding to instruction and whether instruction is responding appropriately to the child. The process for accomplishing these requirements is open, and the available research offers a wide array of options (Scanlon, 2011). However, choices are very much influenced by whether one adopts an identification-measurement or a prevention-instruction frame. Screening is a good example of how perspectives differ.

Screening

Whether one is trying to prevent children from acquiring disabilities or to identify those who already have them, it makes sense to screen students. The RTI component of IDEA says little about screening except that it is necessary for instruction but not central to determining disability. The framers of the law viewed the nature of screening as the domain of states and local districts (Assistance to States for the Education of Children with Disabilities, 2006, p. 46639). In the absence of screening requirements, writers working from a measurement perspective have taken up the language of public health and assumed the need for universal screening (e.g., Jenkins, Hudson, & Johnson, 2007), though this is not specified in the law and not necessarily optimal. There are many approaches to screening, some of which have been obscured by this and related assumptions. For example, one could recognize teachers’ ability to distinguish which children are more or less at risk on the basis of ongoing classroom assessment (Taylor, Anselmo, Foreman, Schatschneider, & Angelopoulos, 2000) and, allowing a buffer for error, test only the lowest half of the students in more detail (Clay, 2004). Indeed, if after a couple of months teachers are unable to determine which students are having more difficulty acquiring literacy than others, it is unlikely that high-quality instruction is taking place, bringing us back to the expertise problem (Sato, Wei, & Darling-Hammond, 2008).

From a measurement perspective, the qualities of screening assessment are sensitivity and specificity. Sensitivity “refers to the accuracy of a screening process in identifying as ‘at risk’ individuals who in fact perform unsatisfactorily on a future criterion measure (i.e., ‘true positives’). Specificity refers to the accuracy of a screening mechanism in identifying as ‘not at risk’ individuals who later perform satisfactorily on a future criterion measure (i.e., ‘true negatives’)” (Jenkins et al., 2007, p.

581). In this frame, high sensitivity and specificity are valued. From a measurement perspective, the ideal is to have an assessment tool that produces no false positives and no false negatives. From an instructional perspective, the ideal is to have an assessment/intervention approach that produces all false positives—that is, everyone predicted to have difficulty using the initial assessment becomes successful because of the subsequent intervention.

In the context of an instructional frame, a screening assessment is simply a tool for deciding where to put resources to maximize prevention and minimize costs. Where the costs are low, specificity does not need to be high and effective screening can be very uncomplicated. For example, Scanlon et al. (2005) found that when children enter kindergarten, a simple test of alphabet knowledge is an excellent predictor of who is likely to encounter difficulty in literacy learning. This simple test allows relatively low-resource prevention efforts, such as focused small-group work, to begin immediately. Some children respond rapidly to such modest interventions, and support can be discontinued on the basis of ongoing classroom assessments, leaving more resources for those still in need. As more data are gathered about writing, reading, and word knowledge, a more detailed examination of routinely collected instructionally relevant data (e.g., writing samples, records of children's reading) might be scheduled (McGill-Franzen et al., 2010).

Where the stakes are high, such as when children will receive intensive one-on-one instruction, we need a sensitive test. The Observation Survey (Clay, 2004) is a good example of such an assessment. It documents the concepts and specific knowledge the child has about print and how it works, plus the strategies he or she uses to make sense with print. It not only screens well but also provides detailed, instructionally useful information (Denton, Ciancio, & Fletcher, 2006). It is very sensitive for a specific stage of development. However, it is noticeably absent from RTI discussions within the measurement framework.

Progress Monitoring

Progress monitoring is a necessary feature of classroom practice and an important element of the RTI legislation. Progress-monitoring assessment strategies are similarly affected by whether one takes up the identification-measurement or prevention-instruction frame. Over the years, teachers and researchers in regular education have developed many productive methods for monitoring progress within an instructional frame. In this frame, the central feature of progress monitoring would be that it offer instructionally useful information as well as indicate progress. For example, in this frame it might make sense to use a measure of instructionally appropriate text level, such as Reading Recovery book levels, for progress monitoring rather than a measure of reading speed on grade-level passages, even if the psychometric properties of the indicator were less adequate. Instructional book levels also make sense because for instruction to be effective, task difficulty is central, not only for the child to be in control of strategic action but also for building task persistence (Gersten, Fuchs, Williams, & Baker, 2001). Indeed, in their meta-analysis of instructional interventions in the learning disabilities field, Swanson and Hoskyn (1998) found that, regardless of the domain of study, managing task difficulty was one of three core instructional variables that explained most of the common variance in outcomes. Instructional book-level data can be collected without requiring children

who are already at risk to waste time reading material that is too difficult for them (and acquiring problematic understandings about reading in the process). Collecting such data would also reveal when children were in fact reading texts that were too difficult—a common and counterproductive fate of less accomplished readers (Allington, 1983).

In fact, many indicators can be used under RTI. For example, just keeping track of a child's word and letter knowledge in kindergarten can be done with one side of a manila folder containing a grid with a cell for each letter of the alphabet and space to note new details about recognition, use in writing, and words known, et cetera. The folder might contain writing samples, running records, and other observations. Regular, brief reading and writing conferences provide updates, and dated samples of writing and running records in the folder provide further evidence and data. Regularly taking stock of these multiple data sources would give a more complete picture of development (McGill-Franzen et al., 2010). Checklists can also be used (Scanlon et al., 2005), particularly if they require supporting evidence.

These measures have been sidelined by the measurement framing of RTI, and CBM has become “the most likely procedure to be used for RTI evaluations of academic performance” (Christ & Hintze, 2007, p. 95). In large part, this is because it was developed in special education and “has become a standard practice in graduate training programs in school psychology” (Olson et al., 2007, p. 120). CBM was developed by Stanley Deno in the late 1970s as a simple assessment practice for special education teachers (Deno, 1985). Initially Deno argued that the speed with which children read the words in their basal textbook was a good indicator of their reading ability because it took little time and used curriculum materials already in use in the classroom; thus such a measure is curriculum based. He also argued that because reading words quickly and accurately required children to integrate all the separate components of reading, the measure would not distort the curriculum by drawing teachers' attention inappropriately to all the individual components (Alt & Samuels, 2011).

Using available curriculum materials and not distorting instruction are both important considerations, particularly for an instructional frame. However, in the intervening years, it was decided that local curriculum materials would not provide the standardization required in a measurement framework, so text passages were brought from outside the classroom and standardized (Fuchs, 2003). This rendered the term “curriculum based” no more appropriate than for any other standardized testing practice. Indeed, the manual of CBM AIMSweb spelling—a series of speeded spelling tests—makes this clear: “Spelling assessment lists are ‘curriculum independent’ allowing teachers to make decisions about general spelling outcomes regardless of spelling curriculum differences between teachers and schools” (Shinn & Shinn, 2002, p. 10). It is misleading to refer to such measures as curriculum based.

Monitoring progress by measuring reading speed and accuracy gained a boost when the report of the National Reading Panel (National Institute of Child Health and Human Development, 2000) popularized the related but different construct of fluency and spawned numerous books on how to teach fluency. The confluence of interest has led to an excessive concern for reading speed. Comprehension has become represented by the number of words said in a given time to retell the story (Good & Kaminski, 2002) or by the number of missing words omitted from a text that the child can accurately replace. The former lacks any construct validity, and the

latter is restricted to a limited view of comprehension at the phrase or sentence level. Indicators of writing development—simply counting changes in the number of words written, for example—are no more promising (Gansle, Noell, VanDerHeyden, Naquin, & Slider, 2002). These assessments focus on the most easily quantified indicators—speed and accuracy at the word level.

There has been considerable criticism of the validity of these measures (Paris, 2005), especially when viewed as indicators of fluency (Samuels, 2007) or reading competence (Alt & Samuels, 2011; Goodman, 2006). Reading speed is not necessarily related to comprehension and is only one aspect of fluency (Cramer & Rosenfield, 2008; Pressley, Hilden, & Shankland, 2005; Valencia et al., 2010). Knowing now that rapid identification of words does not integrate all the components of reading (Samuels, 2007) makes relevant Deno's concern that focusing on partial indicators of reading distracts teachers from the larger goals of literacy instruction (albeit on the very measure he championed). This is a matter of consequential validity to which I will return presently.

CBM is not the only available history of ongoing curriculum-based progress monitoring. For example, running records of children's reading also have an extensive history in progress monitoring and date from the late 1970s (Clay, 1979; Ross, 2004). These are a shorthand record of the words recognized and the strategic processes a child uses while reading. They are widely used in classrooms in the United States and many other countries and have a strong history since the mid-1970s. Vellutino and Scanlon's (2002) successful interventions also have included teacher documentation of the nature of children's errors and strategies and text difficulty in an abbreviated but consistent manner (Scanlon, personal communication, July 3, 2009). These assessment strategies do not appear in writings from special education school psychology because they do not fit the measurement frame. For example, they are not addressed at all in the *Handbook of RTI* (Jimerson et al., 2007b) and are rarely in other RTI writings. Because they document progress and process on materials used in the actual classroom curriculum and are linked directly to instruction, they are more curriculum based than current CBM indicators. Processes can be complex and contextual and require a degree of expertise to recognize and record; it may be that these assessments can be less reliable than some CBM measures. However, the reliability in CBM is gained by focusing on speed and constrained skills (Paris, 2005), and there is a substantial trade-off in instructionally relevant information.

Assessment and Instruction

There are reasons for concern that the CBM "overall indicator approach to progress monitoring has been the more widely used method for gauging student improvement in reading" (Stecker, Fuchs, & Fuchs, 2008, p. 11). First, even if a child only reads for a minute to assess CWPM, recording, organizing, and transitioning takes at least another minute. With 25 children, suddenly an hour is gone. Even once a week adds up to 3.5 hours per month, with no instructionally useful information. This is true even if assessment is done by "swat teams," since personnel could be deployed instructionally instead. Besides, there is little hope of teachers having a sense of ownership of data they did not collect. More importantly, there is the problem that instruction commonly reflects assessment practices, even when the assessment offers no instructionally useful information. At the broad level, for example, if

assessment focuses solely on reading and not writing (often the case), that emphasis is likely to be reflected in instruction, but there are subtler differences that reflect different orientations to teaching and learning.

Referring to CBM, Christ and Hintze (2007) pointed out that “most of the instruments and procedures that are used for RTI evaluations were developed across both behavior analytic and psychometric frameworks” (p. 100), a link that has been documented previously (Shepard, 1991). Measurement-behavior analytic assessment efforts focus on speed and accuracy, making no record of strategic processes such as self-correction. Indeed, self-correction is a prime example of a difference between the behaviorist model championed by school psychology and special education, and a regular education instructional frame. In the assessment process, we determine what it means when a child misreads a word, notices it does not make sense, and goes back and repairs the error. In the context of CBM, its significance, because of the reduced speed, is simply as a negative indicator of development. In an instructional frame, the behavior is an indicator that the child is both monitoring and strategically repairing his reading, an indicator of the development of a self-extending (or executive) system (Clay, 1991; Johnston, 2002; Vellutino & Scanlon, 2002), and thus a positive indicator of development. Self-correction and its significance for independence are not mentioned in the entire *RTI Handbook* (Jimerson et al., 2007b). Indeed, CBM often specifically directs teachers not to value self-correction, particularly by emphasizing speed and time limits. The AIMSweb spelling assessment manual directs teachers to “try to ensure students stop writing as soon as you say ‘stop’ to avoid them going back and adding or correcting” (Shinn & Shinn, 2002, p. 13).

This difference in focus is not trivial, particularly when transferred to instruction. Studies of more and less effective teachers, for example, show that the less effective attend specifically to accuracy rather than strategic processing (Lyons, Pinnell, & DeFord, 1993; Phillips & Smith, 1997), raising real concerns about the consequences of the use of CBM and derivatives such as DIBELS and AIMSweb. Problem solving and self-correcting, which CBM discourages, have the dual benefit of helping children learn to read previously unfamiliar words and keeping them focused on making sense of what they read—comprehending. Records of children’s reading behaviors, because they are process focused, are more likely to lead to strategy-focused instruction that will turn attention away from trait-oriented interpretations of performance and build resilience (Kamins & Dweck, 1999). By focusing solely on speed and accuracy and taking no account of the context of performance, particularly the relative text difficulty, CBM can misdirect teachers’ instructional efforts. This is a serious matter of construct and consequential validity.

CWPM assessment answers the question, how quickly and accurately does this student identify words in grade-level material? often requiring children to read text that is too difficult for them. The information is generalized to the question, how well does this child read? and, by extension, is this child progressing sufficiently slowly to be identified as SLD? By contrast, process-oriented assessments like records of children’s reading processes focus attention on very different questions: Is instruction taking place in appropriately difficult materials? Is the child building a meaning-directed system? Is the child taking control of monitoring and problem solving in reading? How is the child using strategies and resources in the process of making meaning? How is the child’s literacy changing? Each question is instructionally relevant and turns attention toward strategic processing. Validity is about what is val-

ued, and different frames of reference lead to different instructional values and consequences for students' learning.

Research-Based Literacy Instruction and RTI

The RTI component of IDEA is explicitly linked to the NCLB legislation, not only in the common intention to reduce the number of children encountering difficulty in school but particularly in its emphases on reading and research-based instruction. There is no doubt that research has been informative regarding useful dimensions to consider when teaching children to read. The National Reading Panel (National Institute of Child Health and Human Development, 2000) in its original meta-analysis offered both useful considerations and cautions. For example, they point out that they only considered areas in which they felt there was sufficient experimental research, a stance that is not without its problems (Allington, 2006). It has ignored writing and the reciprocal benefits of reading and writing, allowed a view of reading as rapid word recognition, and omitted anything to do with the motivational, relational, or contextual aspects of literacy development. Indeed, recent years have seen frequent examples of poor advice and policy stemming from overgeneralization, consideration of too few factors, overemphasis on the importance of particular factors, and improprieties involving financial and ideological interests (Allington, 2005, 2006; Camilli, Wolfe, & Smith, 2006; Office of the Inspector General United States Department of Education, 2006).

The weaknesses in the research and its uses are becoming apparent. For example, Al Otaiba and Torgesen (2007) reviewed a range of intervention programs deemed successful and argued that, "Given that [state tests] require a much broader range of knowledge and skill than the word-level tests used to estimate success rates . . . it is likely that poor and minority students, in particular, will not achieve the same success rates on them as for the simpler tests that assess only word reading accuracy" (p. 220). Their observation is an important recognition that few interventions currently informing RTI, particularly in the early grades, have addressed much beyond word identification. Indeed, with some notable exceptions (e.g., Dorn & Schubert, 2008; Scanlon et al., 2008), current RTI interventions and assessment appear informed solely by a view of what technical skills need to be known by children in order to read words rather than by a view of how and why children acquire literacy or what makes learning possible for students.

Al Otaiba and Torgesen also noted that the available research is limited in that "we still know little about how best to support the development of vocabulary, conceptual knowledge, reading comprehension, and thinking skills or how to address motivational or behavior management issues" (p. 220). Their observation is, as they intended, a reminder of limited research funding priorities, but it is also an important marker of another problem, our failure to take a broader view of science and look across ideological boundaries for possibilities. Descriptive and analytic research on writing (e.g., Dyson, 1993, 1999, 2002), for example, because it is not experimental research is routinely ignored in RTI publications. There is considerable existing research in each of the areas they cite, such as motivation (e.g., Dweck, 1999; Guthrie et al., 2006). Similarly, there are programs that have been successful in improving the reading ability, including the comprehension, of children experiencing difficulty in literacy. The only one identified on the federal What Works Clearinghouse (WWC)

Web site is Reading Recovery (<http://ies.ed.gov/ncee/wwc/>). Simply considering how this program differs from those that only successfully improve word knowledge could lead to more powerful interventions, but RR receives two dismissive references in the *Handbook of Response to Intervention* (Jimerson et al., 2007b). RR is also institutionally rejected. Even now, while the WWC Web site shows RR's effectiveness, it assiduously avoids mentioning RR when offering advice on selecting RTI programs.

This limited view of literacy and research seems unwise at best and not what is demanded by the law. IDEA's "scientifically based" requirement defers to the definition used in the 2001 No Child Left Behind Act (Title IX—general provisions; part *a*—definitions). The provision defines scientifically based research as "research that involves the application of rigorous, systematic, and objective procedures to obtain reliable and valid knowledge relevant to education activities and programs; and . . . includes research that employs systematic, empirical methods that draw on observation or experiment; [with] . . . rigorous data analyses [and] reliable and valid data . . . across multiple measurements and observations" (sec. 9101. Definitions). The definition notes that studies should be in a peer-reviewed journal or have had the same level of independent rigorous scrutiny by experts; the preference is for randomized control studies with sufficient detail to allow replication. However, the language of the research-based provision also insists that the application of research will be done with due regard to the problems of overgeneralization from groups to individuals (and vice versa), across contexts and populations, and without oversimplifying the complexity of literacy. As Al Otaiba and Torgesen (2007) pointed out in their review of early intervention programs, word-level assessments commonly used to evaluate interventions and that are the heart of CBM will not generalize to the more complex requirements of even state assessments (for this and other problems, see Paris, 2005, 2010), particularly for poor and minority students.

Moving Forward with RTI

RTI offers us an important opportunity to reduce the number of children becoming disabled in literacy, but there is a real danger that its potential will not be realized. Whether through unproductive framing, ideological rigidity, premature closure of possibilities, or efforts to maintain existing institutional roles, RTI is being implemented in ways that are not as productive as they might be. The bottom line in RTI is improved instruction, which requires increasingly expert teachers with instructionally useful data on each child and their own teaching in circumstances in which they can make optimal use of it.

The current identification-measurement approach to RTI may lead to confident classification of children with SLDs in literacy; however, focusing on classification will not necessarily lead to a real reduction in the number of children with reading difficulties. To persist with this path requires a belief not only that a substantial group of children have permanent intrinsic traits (SLDs) that need identification, but that, having identified those traits, we have instruction that specifically addresses the handicap. There is no evidence of this.

There is, however, evidence that we can prevent most children from becoming handicapped in literacy. Some would say that if we accomplished the feat of having all children reading and writing competently we would only have covered up the problem, because some children would still be disabled but just wouldn't look like

it—a problem of “under-identification” (Fuchs & Fuchs, 2006). But McDermott and Varenne (1995) posed the question, “What if the very act of saying there is something wrong, if improperly contextualized, makes [children’s] situation worse?” (p. 139), and offered evidence that it is often the case. Since identification cannot be accomplished directly through assessment, but only as a default interpretation following ensuring appropriate instruction, it seems that the first order of business should be attempting to provide appropriate instruction rather than identifying who possesses a presumed underlying disability. Even if a child becomes classified as SLD, the goal must remain the same—optimizing instruction for that child. The classification provides no instructionally useful diagnostic information, rendering its value tenuous, particularly given the effects of teacher expectation (Rubie-Davies, Hattie, & Hamilton, 2006).

William (2006) observed that “researchers have underestimated the complexity of what it is that teachers do, and in particular, have failed to understand how great an impact context has on teachers’ practice” (p. 1). If our interventions are expected to ensure the best instruction for each student, we cannot avoid attending to the school as a learning institution for teachers as well as students (Dorn & Schubert, 2008). No teacher operates in isolation, and it is common for children needing the most careful and focused instruction to travel among several teachers and specialists, each with their own instructional predilections (Johnston, Allington, & Afflerbach, 1985). When we look for models of RTI, we might consider research showing literate development beyond the level of the word along with the development of teaching capacity and learning community (e.g., Mei Kuin, McNaughton, Amituanai-Tolosa, Turner, & Hsiao, 2009).

If we want to capitalize on the promise of RTI, we must focus on prevention-instruction models, recognizing the complexity of literacy, its teaching and its learning, and centralizing the ongoing development of teacher expertise. None of this can be purchased in canned packages.

Note

This manuscript has been substantially improved by feedback from numerous individuals, including Richard Allington, Donna Scanlon, Frank Vellutino, and Karen Wixson, along with anonymous reviewers from *Educational Researcher*, *Reading Research Quarterly*, and *Elementary School Journal*.

References

- Aaron, P. G. (1997). The impending demise of the discrepancy formula. *Review of Educational Research*, *67*, 461–502.
- Advocacy Institute. (2002). *Students with learning disabilities: A national review*. Burke, VA: Author.
- Al Otaiba, S., & Torgesen, J. (2007). Effects from intensive standardized kindergarten and first-grade interventions for the prevention of reading difficulties. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 212–222). New York: Springer.
- Allington, R. L. (1983). The reading instruction provided readers of differing ability. *Elementary School Journal*, *83*(5), 548–559.
- Allington, R. L. (2005). Ideology is still trumping evidence. *Phi Delta Kappan*, *86*(6), 462–468.

- Allington, R. L. (2006). Reading lessons and federal policy making: An overview and introduction to the special issue. *Elementary School Journal*, *107*(1), 3–15.
- Allington, R. L., & McGill-Franzen, A. (1992). Unintended effects of educational reform in New York State. *Educational Policy*, *6*(4), 396–413.
- Alt, S. J., & Samuels, S. J. (2011). Reading fluency: What is it and how should it be measured? In A. McGill-Franzen & R. Allington (Eds.), *Handbook of reading disability research* (pp. 173–182). New York: Routledge.
- Ardoin, S. P. (2006). The response in response to intervention: Evaluating the utility of assessing maintenance of intervention effects. *Psychology in the Schools*, *43*(6), 713–725.
- Arnold, M., & Lassmann, M. E. (2003). Overrepresentation of minority students in special education. *Education*, *124*(2), 230–236.
- Assistance to States for the Education of Children with Disabilities, 34 C.F.R. § 300 (2006).
- Backes, W., Vuurman, E., Wennkes, R., Spronk, P., Wuisman, M., Engelshoven, J. V., & Jolles, J. (2002). Atypical brain activation of reading processes in children with developmental dyslexia. *Journal of Child Neurology*, *17*(12), 867–871.
- Batsche, G. M., Curtis, M. J., Dorman, C., Castillo, J. M., & Porter, L. J. (2007). The Florida problem-solving/response to intervention model: Implementing a statewide initiative. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 378–395). New York: Springer.
- Batsche, G. M., Elliott, J., Graden, J., Grimes, J., Kovaleski, J. F., & Prasse, D. (2005). *Response to intervention: Policy considerations and implementation*. Alexandria, VA: National Association of State Directors of Special Education.
- Beiser, M., & Gotowicz, A. (2000). Accounting for native/non-native differences in IQ scores. *Psychology in the Schools*, *37*(3), 237.
- Broikou, K. (1992). *Understanding primary grade classroom teachers' special education referral practices* (Unpublished doctoral dissertation). State University of New York at Albany.
- Burns, M. K., & Coolong-Chaffin, M. (2006). Response to intervention: The role of and effect on school psychology. *School Psychology Forum: Research and Practice*, *1*(1), 3–15.
- Callender, W. A. (2007). The Idaho results-based model: Implementing response to intervention statewide. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 331–342). New York: Springer.
- Camilli, G., Wolfe, P. M., & Smith, M. L. (2006). Meta-analysis and reading policy: Perspectives on teaching children to read. *Elementary School Journal*, *107*(1), 27–36.
- Carlson, E., & Parshall, L. (1996). Academic, social and behavioral adjustment for students declassified from special education. *Exceptional Children*, *63*(1), 89–100.
- Catts, H. W., Hogan, T. P., & Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities*, *36*(2), 151.
- Christ, T. J., & Hintze, J. M. (2007). Psychometric considerations when evaluating response to intervention. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 93–105). New York: Springer.
- Clay, M. M. (1979). *Reading: The patterning of complex behaviour* (2nd ed.). Auckland, New Zealand: Heinemann.
- Clay, M. M. (1985). *The early detection of reading difficulties: A diagnostic survey with recovery procedures* (3rd ed.). Auckland, New Zealand: Heinemann.
- Clay, M. M. (1990). The reading recovery programme, 1984–88. *New Zealand Journal of Educational Studies*, *25*(1), 61–70.
- Clay, M. M. (1991). *Becoming literate: The construction of inner control*. Portsmouth, NH: Heinemann.
- Clay, M. M. (2004). *An observation survey of early literacy achievement* (2nd ed.). Portsmouth, NH: Heinemann.
- Clay, M. M. (2005). *Literacy lessons designed for individuals: Part two, teaching procedures*. Portsmouth, NH: Heinemann.
- Cramer, K., & Rosenfield, S. (2008). Effect of degree of challenge on reading performance. *Reading and Writing Quarterly*, *24*(1), 119–137.

- Darling-Hammond, L., & McLaughlin, M. W. (1999). Investing in teaching as a learning profession: Policy problems and prospects. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 376–411). San Francisco: Jossey-Bass.
- Deno, S. (1985). Curriculum based measurement: The emerging alternative. *Exceptional Children*, *52*(3), 219–232.
- Denton, C. A., Ciancio, D. J., & Fletcher, J. M. (2006). Validity, reliability, and utility of the Observation Survey of Early Literacy Achievement. *Reading Research Quarterly*, *41*(1), 8–34.
- Deshler, D. A. (2008, July 1). *Fidelity! Fidelity! Fidelity!—What About Dosage?* Retrieved from <http://www.rtinetwork.org/rti-blog/entry/1/12>
- Dorn, L. J., & Schubert, B. (2008, Spring). A comprehensive intervention model for preventing reading failure: A response to intervention process. *Journal of Reading Recovery*, *7*(2), 29–41.
- Drame, E. R. (2002). Sociocultural context effects on teachers' readiness to refer for learning disabilities. *Exceptional Children*, *69*, 41–53.
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology Press.
- Dyson, A. H. (1993). *Social worlds of children learning to write in an urban primary school*. New York: Teachers College Press.
- Dyson, A. H. (1999). Coach Bombay's kids learn to write: Children's appropriation of media material for school literacy. *Research in the Teaching of English*, *33*(4), 367–402.
- Dyson, A. H. (2002). *The brothers and sisters learn to write*. New York: Teachers College Press.
- Elementary and Middle Schools Technical Assistance Center. (2009). *Disproportionality: The Disproportionate Representation of Racial and Ethnic Minorities in Special Education*. Retrieved from <http://www.emstac.org/registered/topics/disproportionality/intro.htm>
- Fletcher, J. M. (1992). The validity of distinguishing children with language and learning disabilities according to discrepancies with IQ: Introduction to the special series. *Journal of Learning Disabilities*, *25*(9), 546–548.
- Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, *41*(1), 92–99.
- Fuchs, D., Stecker, P. M., & Fuchs, L. S. (2008). Tier 3: Why special education must be the most intensive tier in a standards-driven, No Child Left Behind world. In D. Fuchs, L. S. Fuchs, & S. Vaughn (Eds.), *Response to intervention: A framework for reading educators* (pp. 71–104). Newark, DE: International Reading Association.
- Fuchs, L. S. (2003). Assessing intervention responsiveness: Conceptual and technical issues. *Learning Disabilities Research & Practice*, *18*(3), 172–186.
- Gansle, K. A., Noell, G. H., VanDerHeyden, A. M., Naquin, G. M., & Slider, N. J. (2002). Moving beyond total words written: The reliability, criterion validity, and time cost of alternative measures for curriculum-based measurement in writing. *School Psychology Review*, *31*(4), 477–497.
- Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research*, *71*(2), 279–320.
- Gilmore, A. (2002). Large-scale assessment and teachers' assessment capacity: Learning opportunities for teachers in the National Education Monitoring Project in New Zealand. *Assessment in Education: Principles, Policy & Practice*, *9*(3), 343–361.
- Good, R. H. I., & Kaminski, R. A. (2002). *Dynamic indicators of basic early literacy skills* (Rev. ed., July 2, 2003). Retrieved from https://dibels.uoregon.edu/measures/files/admin_and_scoring_6th_ed.pdf
- Goodman, K. (2006). *The truth about DIBELS: What it is, what it does*. Portsmouth, NH: Heinemann.
- Gresham, F. M. (2007). Evolution of the response-to-intervention concept: Empirical foundations and recent developments. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 10–24). New York: Springer.
- Guthrie, J. T., Wigfield, A., Humenick, N. M., Perencevich, K. C., Taboada, A., & Barbosa, P. (2006). Influences of stimulating tasks on reading motivation and comprehension. *Journal of Educational Research*, *99*(4), 232–245.

- Harold, D., Paracchini, S., Scerri, T., Dennis, M., Cope, N., Hill, G., . . . Monaco, A. P. (2006). Further evidence that the KIAA0319 gene confers susceptibility to developmental dyslexia. *Molecular Psychiatry*, *11*(12), 1085–1091.
- Heath, S. B. (1983). *Ways with words*. Cambridge, MA: Cambridge University Press.
- Hoefl, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J. L., . . . Gabrieli, J. D. E. (2007). Functional and morphometric brain dissociation between dyslexia and reading ability. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(10), 4234–4239.
- Jakupcak, M. J. (1975). Areas of congruence in remedial reading and learning disabilities. *Journal of Special Education*, *9*(2), 155–157.
- Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for at-risk readers in a response to intervention framework. *School Psychology Review*, *36*(4), 582–600.
- Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (2007a). Response to intervention at school: The science and practice of assessment and intervention. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 3–9). New York: Springer.
- Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (Eds.). (2007b). *Handbook of response to intervention: The science and practice of assessment and intervention*. New York: Springer.
- Johnston, P. H. (2002). Commentary on “The interactive strategies approach to reading intervention.” *Contemporary Educational Psychology*, *27*, 636–647.
- Johnston, P. H., Allington, R. L., & Afflerbach, P. (1985). The congruence of classroom and remedial reading programs. *Elementary School Journal*, *85*(4), 465–478.
- Kamins, M. L., & Dweck, C. S. (1999). Person versus process praise and criticism: Implications for contingent self-worth and coping. *Developmental Psychology*, *35*(3), 835–847.
- Kavale, K. A. (1990). Effectiveness of special education. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (pp. 868–898). New York: Wiley.
- Knotek, S. E. (2007). Consultation within response to intervention models. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 53–64). New York: Springer.
- Kovaleski, J. F. (2007). Response to intervention: Considerations for research and systems change. *School Psychology Review*, *36*, 638–646.
- Lyon, G. R., Fletcher, J. M., Shaywitz, S. E., Shaywitz, B. A., Torgesen, J. K., Wood, F. B., . . . Olson, R. (2001). Rethinking learning disabilities. In J. C. E. Finn, R. A. J. Rotherham, & J. C. R. Hokanson (Eds.), *Rethinking special education for a new century* (pp. 259–287). Washington, DC: Progressive Policy Institute and the Thomas B. Fordham Foundation.
- Lyons, C. A., Pinnell, G. S., & DeFord, D. E. (1993). *Partners in learning: Teachers and children in Reading Recovery*. New York: Teachers College Press.
- McDermott, R., & Varenne, H. (1995). Culture as disability. *Anthropology and Education Quarterly*, *26*(3), 324–348.
- McGill-Franzen, A., Allington, R. L., Yokoi, L., & Brooks, G. (1999). Putting books in the classroom seems necessary but not sufficient. *Journal of Educational Research*, *93*(2), 67–74.
- McGill-Franzen, A., Payne, R., & Dennis, D. (2010). Responsive intervention: What is the role of appropriate assessment? In P. H. Johnston (Ed.), *RTI in literacy—Responsive and comprehensive* (pp. 115–132). Newark, DE: International Reading Association.
- Mei Kuin, L., McNaughton, S., Amituanai-Tolosa, M., Turner, R., & Hsiao, S. (2009). Sustained acceleration of achievement in reading comprehension: The New Zealand experience. *Reading Research Quarterly*, *44*(1), 30–56.
- National Association of School Psychologists. (2000). *Standards for training and field placement programs in school psychology and standards for the credentialing of school psychologists*. Retrieved from <http://www.nasponline.org/standards/FinalStandards.pdf>
- National Center for Learning Disabilities. (2008). National online forum video Webcast, June 9, 2008. Retrieved from <http://www.connectlive.com/events/rtnetwork060908/rtnetwork060908-archive.aspx>
- National Center for Special Education Research. (2007). *Perceptions and expectations of youth with disabilities (NLTS2)*. Retrieved from http://ies.ed.gov/ncser/pubs/20073006/tables/table_a5.asp
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research liter-*

- ature on reading and its implications for reading instruction (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- Nelson, J. M., & Machek, G. R. (2007). A survey of training, practice, and competence in reading assessment and intervention. *School Psychology Review*, *36*(2), 311–327.
- Office of the Inspector General U.S. Department of Education. (2006). *The Reading First program's grant application process final inspection report*. In U.S. Department of Education (Ed.) (Vol. ED-OIG/I13-F0017).
- Olson, S. C., Daly, E. J., Anderson, M., Turner, A., & LeClair, C. (2007). Assessing student response to intervention. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention*. New York: Springer.
- Pancsofar, N., & Vernon-Feagans, L. (2006). Mother and father language input to young children: Contributions to later language development. *Journal of Applied Developmental Psychology*, *27*(6), 571–587.
- Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, *40*(2), 184–202.
- Paris, S. G. (2010). Thinking straight about measures of early reading development. In P. H. Johnston (Ed.), *RTI in literacy—Responsive and comprehensive* (pp. 103–114). Newark, DE: International Reading Association.
- Phillips, G., & Smith, P. (1997). *A third chance to learn: The development and evaluation of specialized interventions for young children experiencing difficulty with learning to read* (No. 13227). Wellington, New Zealand: National Council for Educational Research.
- Pressley, M., Hilden, K., & Shankland, R. (2005). *An evaluation of end-grade-3 Dynamic Indicators of Basic Early Literacy Skills (DIBELS): Speed reading without comprehension, predicting little*. East Lansing: Literacy Achievement Research Center, Michigan State University.
- Reschly, D. J., & Wilson, M. S. (1995). School psychology practitioners and faculty: 1986 to 1991–92 trends in demographics, roles. *School Psychology Review*, *24*(1), 62–80.
- Roehrig, A. D., Duggar, S. W., Moats, L., Glover, M., & Mincey, B. (2008). When teachers work to use progress monitoring: Data to inform literacy instruction identifying potential supports and challenges. *Remedial & Special Education*, *29*(6), 364–382.
- Ross, J. A. (2004). Effects of running records assessment on early literacy achievement. *Journal of Educational Research*, *97*(4), 186–194.
- Rubie-Davies, C., Hattie, J., & Hamilton, R. (2006). Expecting the best for students: Teacher expectations and academic outcomes. *British Journal of Educational Psychology*, *76*(3), 429–444.
- Samuels, S. J. (2007). The DIBELS tests: Is speed of barking at print what we mean by reading fluency? *Reading Research Quarterly*, *42*(4), 568–570.
- Sato, M., Wei, R. C., & Darling-Hammond, L. (2008). Improving teachers' assessment practices through professional development: The case of National Board Certification. *American Educational Research Journal*, *45*(3), 669–700.
- Scanlon, D. M. (2011). Response to intervention as an assessment approach. In A. McGill-Franzen & R. L. Allington (Eds.), *Handbook of reading disabilities research* (pp. 139–148). New York: Routledge.
- Scanlon, D. M., Gelzheiser, L. M., Vellutino, F. R., Schatschneider, C., & Sweeney, J. M. (2008). Reducing the incidence of early reading difficulties: Professional development for classroom teachers versus direct interventions for children. *Learning & Individual Differences*, *18*(3), 346–359.
- Scanlon, D. M., & Vellutino, F. R. (1996). Prerequisite skills, early instruction and success in first-grade reading: Selected results from a longitudinal study. *Mental Retardation and Developmental Disabilities Research Reviews*, *2*, 54–63.
- Scanlon, D. M., & Vellutino, F. R. (1997). A comparison of the instructional backgrounds and cognitive profiles of poor, average, and good readers who were initially identified as at risk for reading failure. *Scientific Studies of Reading*, *1*(3), 191–215.
- Scanlon, D. M., Vellutino, F. R., Small, S. G., Fanuele, D. P., & Sweeney, J. M. (2005). Severe reading difficulties, can they be prevented? A comparison of prevention and intervention approaches. *Exceptionality*, *13*(4), 209–227.
- Shepard, L. A. (1991). Psychometricians' beliefs about learning. *Educational Researcher*, *20*(7), 2–16.

- Shinn, M. R. (2002). Best practices in using curriculum-based measurement in a problem solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (Vol. 1, pp. 671–697). Bethesda, MD: National Association of School Psychologists.
- Shinn, M. R., & Shinn, M. M. (2002). *AIMSweb training workbook: Administration and scoring of spelling curriculum-based measurement (S-CBM) for use in general outcome measurement*. Retrieved from http://www.aimsweb.com/uploads/pdfs/scoring_scbm.pdf
- Sleeter, C. (1986). Learning disabilities: The social construction of a special education category. *Exceptional Children*, *53*(1), 46–54.
- Stanovich, K. (1991). Discrepancy definitions of reading disability: Has intelligence led us astray? *Reading Research Quarterly*, *26*, 7–29.
- Stecker, P. M., Fuchs, D., & Fuchs, L. S. (2008). Progress monitoring as essential practice within response to intervention. *Rural Special Education Quarterly*, *27*(4), 10–17.
- Swanson, H. L., & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research*, *68*, 277–321.
- Taylor, B. M., Pearson, P. D., Clark, K., & Walpole, S. (2000). Effective schools and accomplished teachers: Lessons about primary grade reading instruction in low income schools. *Elementary School Journal*, *101*(2), 121–165.
- Taylor, H. G., Anselmo, M., Foreman, A. L., Schatschneider, C., & Angelopoulos, J. (2000). Utility of kindergarten teacher judgments in identifying early learning problems. *Journal of Learning Disabilities*, *33*(2), 200–210.
- Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research and Practice*, *15*(1), 55–64.
- U.S. Department of Education. (2002). *The facts: Reading achievement. No Child Left Behind*. Retrieved from <http://www.nochildleftbehind.gov/start/facts/reading.html>
- Valencia, S. W., Smith, A. T., Reece, A. M., Li, M., Wixson, K. K., & Newman, H. (2010). Oral reading fluency assessment: Issues of construct, criterion, and consequential validity. *Reading Research Quarterly*, *45*(3), 270–291.
- Vaughn, S., Levy, S., Coleman, M., & Bos, C. S. (2002). Reading instruction for students with LD and EBD: A synthesis of observation studies. *Journal of Special Education*, *36*(1), 2–13.
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology & Psychiatry*, *45*(1), 2–40.
- Vellutino, F. R., & Scanlon, D. M. (2002). The interactive strategies approach to reading intervention. *Contemporary Educational Psychology*, *27*(4), 573–635.
- Vellutino, F. R., Scanlon, D. M., & Lyon, G. R. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers. *Journal of Learning Disabilities*, *33*(3), 223–238.
- Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, *88*, 601–638.
- Vellutino, F. R., Scanlon, D. M., Small, S., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities*, *39*(2), 157–169.
- Wanzek, J., & Vaughn, S. (2008). Response to varying amounts of time in reading intervention for students with low response to intervention. *Journal of Learning Disabilities*, *41*(2), 126–142.
- Westat. (2009). *28th annual report to Congress on the implementation of the Individuals with Disabilities Education Act, 2006*. Retrieved from <http://www2.ed.gov/about/reports/annual/osep/2006/parts-b-c/28th-vol-1.pdf>
- William, D. (2006). *Does assessment hinder learning?* Address presented at the Educational Testing Service's Europe Breakfast Seminar: Educational Testing Services. Retrieved from http://www.uk.etsEurope.org/home-corpo-uk/news-home/?print=1&news=136&view=detail&no_cache=1