

# **To DIBELS or Not to DIBELS?**

## **Answering Questions about Fluency-Based Measures in the Era of Response to Intervention**

By

Michael P. Ford

University of Wisconsin Oshkosh

*“DIBELS is the worst thing to happen to the teaching of reading  
since the development of flashcards.”*

*P. David Pearson*

*“Anyone using DIBELS to make important decisions about students’ development of reading  
proficiencies is guilty of educational malpractice.”*

*Richard Allington*

[Author’s note: This article has been in the publication pipeline for a long time. When first prepared, it focused on the external critique of a specific assessment tool in its original implementation forms and formats. Since that time the tool itself has under gone changes and now after sixth editions, *DIBELS Next* has emerged. Its website suggests seven changes: all new forms and passages; improved, clearer directions for assessor and students; new easier-to-administer measure of early phonemic awareness with First Sound Fluency replacing Initial

Sound Fluency; new measure based on maze procedures, Daze, added as an additional indicator of comprehension for third through sixth grade; larger format for scoring booklets to facilitate scoring and the recording of response patterns; retelling integrated into the Oral Reading Fluency measure; and new benchmark goals and cut points for risk (<http://dibels.org/next.html>.) It seems that every time one thinks fluency-based measures are finally fading from current practices, they resurface. Now as debates on screeners for the primary grades receive political attention in this state, it may be time to revisit the original concerns about assessment tools like DIBELS and its many variations.

That leads to a second important change since the article was first drafted. There are now a number of fluency-based measures used as CBMs for progress monitoring in RtI frameworks. As one WSRA member observed, “because of the concerns that have been raised about DIBELS specifically in the last few years, many districts are leaning more toward other similar systems such as Easy CBMs and AIMSweb...” So while this article looks specifically at one tool, the concerns the article raises are often true of other fluency-based measures used in RtI frameworks.]

### **What is DIBELS?**

DIBELS is an acronym that stands for Dynamic Indicators of Basic Early Literacy Skills. While recently revised to include grades 3-6, the authors of its original version claimed that it was a standardized set of seven individually administered measures of early literacy development usually given on a wide-scale basis to students in kindergarten and first grade. They were labeled *fluency* measures and designed to be short (one minute) measures used to regularly monitor development of prereading and early reading skills. They included the following subtests: Letter

Naming Fluency, Initial Sounds Fluency (revised as First Sound Fluency), Phoneme Segmentation Fluency, Nonsense Word Fluency, Oral Reading Fluency, Retelling Fluency (now integrated with Oral reading Fluency) and Word Use Fluency. [Consult Peter Afflerbach's book *Understanding and Using Reading Assessment K-12* for a good independent description and discussion of DIBELS especially as contrasted with another popular early literacy assessment tool Clay's An Observation of Early Literacy Achievement (OSELA).]

### **What is the purpose of fluency-based measures?**

Fluency-based measures have been adopted by many districts as one way –sometimes the only way -- to assess whether a kindergarten or first grade aged child has been successful in responding to classroom reading instruction. It often serves as an initial universal screening tool and then is subsequently used to monitor student progress. The need for this type of curriculum-based measurement (CBM) with progress monitoring has intensified due to changes in federal policies regarding special education referral and placement procedures especially for students with learning disabilities (Samuels, 2005). Recent changes in federal policies specifically in the 2004 reauthorization of the Individuals with Disabilities Act (IDEA) have shifted identification of students with these needs away from requiring the exclusive use of discrepancy models. In these models, a student would have to show a two-year discrepancy or gap between his/her measured IQ and his/her classroom achievement in order to qualify for special education services. This often meant that young children who had not been in school long enough to show this two-year gap could not qualify for services as early as they might be needed. Discrepancy models often were criticized for being too reactive. The changes have led to the recommended use of Response to Intervention (RtI) models.

### **What are Response to Intervention (RtI) models?**

IDEA's reference to RtI and the Department of Education's promotion of the practice has placed these models at the center of much discussion in local schools. RtI initiatives recommend that the identification of students with learning disabilities uses evidence that the students have been unable to successfully respond to instruction/interventions prior to being referred and placed in special education services. In the past, the discrepancy model often allowed for identification and placement regardless of the quality of instruction and/or interventions provided that student.

While best practices seen in many school districts would have included documented attempts to provide students adequate instruction and intervention before placement in special education services, the current changes have given RtI models a more prominent role at the heart of the referral, identification and placement processes. RtI models are often seen as more proactive allowing for identification as soon as the child is not responding to classroom instruction (though it should be noted that this lack of response does not mean that special education services will immediately follow.)

### **That seems like an important positive shift, so why is there so much concern about this view of RtI?**

This view of RtI is described by some as a "treatment resistor" model. Problems are only identified when a student doesn't respond to ("resists") the instruction ("treatment") and then the *problem* is primarily the student's. For example, look at this definition of RtI (Burggraf & Sotomayor, 2007): "The RtI model *assumes* [emphasis added] that all students receive research-based, high-quality, differentiated instruction from a general educator in a general education

setting.” And herein lies a major problem, for it is based on the assumption that all students have had access to quality instruction to begin with but the progress monitoring within CBM does very little to assess the quality of instruction a student has received. It places the emphasis on what the learner is doing in response to instruction and virtually no attention to the quality of instruction the learner is responding to in the first place. In fact in a recent issue of the WSRA Journal, Connecticut’s framework was presented (Costello, 2008). Amazingly enough in providing a rationale for why the framework was needed, they cited six data-based performance trends. Every trend was student-focused. They acknowledged that observations of classroom educators and analysis of student study teams reports led them to conclude that the repertoire of general educators’ instructional strategies were too limited to meet the needs of all the students in their classrooms. They also concluded that the ability to differentiate instruction was not evident. But in presenting Connecticut’s Three Tier Model, instructional controls seem less evident than student controls. In Tier One, the school and district must ensure a number of key instructional components ranging from ensuring a K-12 curriculum to supportive school leadership. The struggles of a learner, however, are always the catalyst for review. While the review is supposed to include a look at strategies used and student response to them, one knows that issues related to other elements supposedly ensured at the first tier -- the K-12 curriculum, ongoing professional development, existence of professional learning communities, ample materials and supportive leadership – will receive very little examination. This is certainly true with the way some systems are used sometimes turning attention to the learner and away from the instruction.

The bottom line is we need to think about these scenarios. If you can guarantee that a student has received high quality instruction, then developing a monitoring system that identifies

a student's inability to make progress is probably effective in red flagging potential problems for the student. If you can't guarantee that a student has received high quality instruction, however, what is the monitoring system really monitoring? The student's inability to make progress may have less to do with the student and more to do with the instruction. In fact, making progress in poor quality instruction could actually give the illusion of appropriate student response when it may simply mean that a student with real problems is merely doing inappropriate tasks well.

### **How could that happen?**

Cambourne (2001) reminds us of six reasons why some children fail to learn to read. First, the students may get faulty demonstrations of how to read and write. So whose fault is it if the student is not learning because the demonstrations they received were either incorrect or inadequate? How does a student-centered progress monitoring system indicate whether a child is receiving faulty demonstrations? Second, the student may receive quality demonstrations but not engage with them. Engagement involves a combination of feeling that one can be successful with the instructional task, values the instructional outcome and feels safe within the instructional environment. Again progress monitoring systems do very little to determine if engagement issues were at the root of a student's struggles and if they were, could they tell which aspect of engagement needs to be addressed? Third, the majority of struggling readers have low expectations of themselves as readers and writers. They no longer see themselves as inside players and this also negatively impacts their levels of engagement. These low expectations are often linked to previous classroom practices that have labeled and sorted children. But again, progress monitoring does not address issues of expectations. It starts with the assumption that all students have been held to high expectations. Fourth, the struggling reader relies on feedback to

grow stronger. But how is the quality of feedback being given students monitored? It's not. Like demonstrations, the quality of feedback can only be assessed by evaluating instruction not an responsibility for their learning. You may be able to use progress monitoring to determine that happened, but not the reasons why. Finally, Cambourne reminds us that a student may struggle for any combination of the five reasons. What becomes very clear in Cambourne's analysis is that the reasons for less successful readers often lie in the instruction they are receiving, yet learner-focused monitoring systems often start with the assumption that all students have received quality instruction.

One problem may be even worse than not monitoring instruction. If a progress monitoring system within CBM assesses something that is far removed from the more global acts of reading and writing (say for example how fast a child can say nonsense words), and a teacher refocuses instruction on doing the assessment task well versus what leads to more effective reading and writing; the student may become expert at saying nonsense words fast while still struggling with the more global demands of reading and writing (Opitz, Ford & Erikson, 2011). The monitoring device gives the illusion that everyone is on track, but actually without attending to what is being taught covers up the real cause of problems students may subsequently reveal. Consider the findings of Walczyk and Griffith-Ross (2007). They reported that the fix-up strategies competent readers use to deal with comprehension difficulties (ie, slowing down and rereading) are actually penalized by systems which focus exclusively on speed. Scoring well in these systems can mask whether early readers are developing and using fix-up strategies they need to deal with comprehension difficulties. This may be one reason why maze-like measures to monitor comprehension have been added to some of these assessment systems.

## **Even with this flaw, wouldn't treatment resistor models and curriculum-based measurement have value?**

Treatment resistor models and CBM can have value if educators first look for patterns across students' responses. In other words if the assessment task is authentic, and it reveals that a lot of students seem to be having difficulty with some dimension of reading development, this pattern probably says more about what needs to be adjusted in the instruction. Patterns should be seen first as potential indicators of need to improve instruction instead of automatic indicators of students' failure to learn. But in many current conceptualizations of RtI, the reaction to students' lack of progress is often to label and sort the students. In popular tiered models of RtI, the students are moved to different more intensive treatments, while the poor quality instruction the students received in the first place remains in place. Most RtI models place the burden on the students and add new layers of instruction and programs instead of guaranteeing students quality instruction in the first place.

## **Why was the change to RtI models made?**

One can speculate about why this change occurred. As we have seen in the past, clearly what looks like on the surface an important pedagogical shift is often the result of intentions that may have more to do with politics and economics. Even with its flaws and when other motives seem less noble, the basic premise of RtI models should not be lost. It is important that any child who is being identified for and placed in special education services has had access to quality instruction and intervention before that decision is made. RtI models when conceived in their best forms have the potential to do that. This may help address a growing concern about the over identification and placement of children in special education services particularly in the area of



learning disabilities. Beyond the potential impact that would have on individual students, obviously a reduction of those identified and placed in special education services has a significant economic impact on local districts. This potential impact has caught the eye of certain policymakers as well. A reduced economic need in this area either allows for a reallocation of funds to other areas (of greater interest to some advocates) or a general reduction of public funding to schools (a motive for some other advocates.)

### **So what role do fluency-based measures play in RtI models?**

Since RtI models are often described with initial universal baseline assessments and significant progress monitoring systems within curriculum-based measurement (CBM), many districts are looking for such systems. In the best practices of many school districts, assessment drives instruction. If educators are expected to determine whether children are successfully responding to quality instruction presented in a classroom reading and/or school intervention program, ongoing systematic progress monitoring is key. Fluency-based measures present themselves as tools for meeting that need.

### **Are fluency-based measures the only acceptable tools for meeting the monitoring expectations of RtI models?**

Of course not (unless it becomes mandated through state policy.) Many school districts already had in place effective systems for monitoring students' progress as a requisite step in providing quality classroom reading instruction and/or school intervention programs. Many had already developed internal CBM systems more closely aligned with local standards, benchmarks and expectations; and more compatible with local philosophies, methods and materials. Other

districts invested in external assessment systems that relied more closely on process-oriented authentic assessment tasks like reading benchmark leveled texts and/or writing samples. There are even other skills-based assessment systems like DIBELS that have been chosen by districts. The problem is that many times DIBELS is presented as the only or best way to satisfy progress monitoring requirements. For example, the popular guide **RTI: A Practitioner's Guide to Implementing Response to Intervention**, Mellard and Johnson (2008) only identify three computer-based examples of CBM including DIBELS. In another early guide **Response to Intervention: A Practical Guide for Every Teacher**, Bender and Shores (2007) present a standard protocol RtI model that features the exclusive use of DIBELS as its curricular probes. Given these two popular guides, DIBELS was perceived as if not the only way, at least one of the best ways, to do progress monitoring in a relatively short amount of time. Afflerbach (2007) cautions however: "...it appears that the adoption of DIBELS may be influenced by more than the dynamic of educated consumers choosing assessments that best suit their needs."

Ironically many schools are now reconciling their initial frameworks for RtI with their embrace of Common Core State Standards and looming assessments aligned with those standards. As Wixson and Lipson (2012) note, "RtI assessment often rests on gathering data for screening and progress monitoring using measures that focus on fluency. Given the CCSS conceptualization of ELA, an emphasis on fluency will barely scratch the surface of either the areas of assessment or the types of measures needed." Districts who abandoned more authentic standards-driven assessment systems due to RtI pressures, now find themselves rethinking those decisions.

### **So how did fluency-based measures become the most pervasive tools for meeting the monitoring demands of RtI models?**

In his analysis, David Pearson (2006) suggests five factors in the popular use of DIBELS. It had a perceived simplicity and ease of use, a transparent alignment with NCLB, was marketed with other commercial materials, created a sense of scientific cachet, and was politically positioned. In other words, DIBELS' popularity was in part due to timing and in part due to effective marketing. Others such as Manzo (2006) and Allington (2009) have suggested that the popularity of DIBELS grew more from its political positioning rather than the quality of the tool. Whatever the reason, DIBELS seemed to be designed and disseminated deliberately to easily meet a growing perceived need that many schools had. It's not surprising so many schools latched on to it. Many of these schools saw a tool that could be downloaded for free. It appeared easy to administer by virtually anyone with little training. The tests were short requiring a minimal amount of time. Those who decided to use it saw the tool as addressing key elements of NCLB on which they were already focused. They also found a tool that could fit easily with some of the commercial materials they were already using.

### **So what's wrong with fluency-based measures?**

The major concern about tools like these is how they create the appearance of scientific cachet. Most sell themselves not on the basis of what they actually assesses, but on the correlation or predictive value between what they assess and more significant dimensions of the reading process. They often never assess more global dimensions of reading (e.g., comprehension strategies, metacognition). They assess "indicators (e.g., letter naming, word calling)." This type of small parts assessment of indicators removed from the global dimensions of reading or overall

reading performance is always problematic because it fails to capture what thoughtful reading requires. Tools that do capture thoughtful reading must demonstrate that they are effectively assessing multiple indicators and then that they have to do with critical literacy skills and behaviors.

Goodman and others (2006) pointed out that initially DIBELS fell short on meeting this first criterion. There were many inherent problems with the original tasks DIBELS used and questions about whether they really indicated critical early literacy skills. But even if that wasn't a problem, DIBELS must demonstrate that these early literacy skills have any connection to overall reading performance and achievement. Instead of developing an assessment tool or process that truly captured the essence of the reading process, DIBELS used what Pearson (2006) called "psychometric alchemy" to suggest that is what the test is doing. In other words, a deeper look at the way some tools correlate measures with global dimensions of reading also revealed the use of less than "gold standard" measures for those correlations. Many tools do not even bother to correlate themselves with what reading experts would label as the best forms of standardized literacy assessments. Experts have concluded that the way some link their tasks to predictive values of overall reading performance is more hocus-pocus than accepted psychometric practices. [It is interesting the one new subtest in DIBELS Next is a maze-like assessment for assessing a more global dimension of reading –comprehension-- in grades 3-6.]

### **How does this statistical manipulation cause problems?**

Using the work of Carlisle, Schilling, Scott and Zeng (2004), Michael Pressley pointed out the initial problems with DIBELS as a sole indicator of reading performance. As he explained, doing so has both the possibility of producing false positives (identifying children at risk when they are

not as indicated by more global measures) and false negatives (not identifying children at risk when they clearly are). In his analysis, he made clear that in one fairly large sample of over 1700 students, DIBELS missed identifying reading problems in one out of every six students. Pressley concluded: "...DIBELS mis-predicts reading performance on other assessments much of the time and at best is a measure of who reads quickly without regard to whether the reader comprehends what is read." This mis-prediction shows how problematic it can be to rely on a system that relies on measuring the small components of reading to monitor the progress of more global dimensions.

Pressley's concerns raise the question of why in an era of so-called scientifically-based research, DIBELS' pervasive use preceded any independent level of review? Likewise, Allington (2009) has identified a number of independent reviews with an overall conclusion that "DIBELS does not measure reading rate and accuracy with any reliability" which continued to call into question its ability to use these measures to accurately predict reading comprehension.

### **Doesn't any subtest have predictive value?**

In one contrary study of over 1500 first grade students by Riedl (2007), the DIBELS Oral Reading Fluency (ORF) was the best predictor of comprehension at the end of first or second grade. (It was actually a better predictor than the Retelling Fluency subtest which according to the developers, is supposed to be the comprehension measure for DIBELS. These now have been combined in DIBELS Next.) ORF successfully predicted comprehension success at the end of first grade for 80% of the students and at the end of second grade for 71%. Further analysis of these statistics, however caused Riedl to conclude that the minimal gains in predictive accuracy using any of the other subtests separately or in combination with the ORF did not justify the time

and energy needed to administer those tests. In other words, if any of the DIBELS subtests have predictive value it is probably the ORF. Important to note, however, is that even in Reidl's analysis misidentification happened for part of the sample. Reidl's analysis revealed a student's level of vocabulary had a compounding factor. Students could do well on ORF, but if they had limited vocabulary, they still had difficulties with comprehension (false positives). On the other hand students who didn't do well on ORF but did adequately on comprehension measures often had higher levels of vocabulary (false negatives). Unfortunately, DIBELS doesn't address vocabulary levels, so the educator must interpret the results of the ORF in light of the student's level of vocabulary using some additional measure.

This interpretation of ORF seems to be echoed in a review by Goffreda and DiPerna (2010) who concluded:

educational professionals use DIBELS to identify students who are in need of early intervention. The purpose of this review was to synthesize the current psychometric evidence for each DIBELS indicator. Strong reliability and validity evidence was observed for DIBELS Oral Reading Fluency; however, evidence for the remaining DIBELS indicators demonstrated greater variability. Although the majority of evidence focused on individual score reliability and validity for single-point decisions, further studies are needed to determine effective practices for progress monitoring.

**How does a tool misuse psychometrics to present itself as a psychometrically sound assessment tool?**

Some fluency-based measures fail to recognize the difference between what reading experts call “constrained skills” and “unconstrained skills.” Constrained skills are an aspect of the reading process that can be mastered like knowing the names of letters or the sounds letters make. In other words, the key assessment task is to check to see whether the child knows this skill or not. A tool that introduces an additional element of speed to this skill is giving value to how fast a student can say the letters. Practically, however, it is hard to point to the significance of being able to name one’s letters in 26 seconds or 25 seconds; but many tools treat those students as different. By introducing the element of speed (or what they label as *fluency*) to these constrained tasks, tools can create an artificial degree of variability in the learners. Then when attempting to correlate those results with an unconstrained skill (comprehension) measured by more global reading assessments, they can artificially create a higher degree of correlation. It is easier to demonstrate reliability and validity with artificially-created variability. Doing so creates an illusion that these small part assessments are really “indicators” of something more global. This is what Pearson means when he talks about psychometric alchemy. Imagine instead if students were just ranked according to whether they had mastered the task or not. This type of ranking creates less variability among those students and makes it harder to show that mastery of a small constrained skill is a reliable indicator of something more important. [Readers are reminded to revisit the article by Scott Paris reprinted in the RtI at a Crossroads: Part One “Reinterpreting the Development of Reading Skills.” Paris concluded:

Educators should be wary of policies that require repeated assessments of constrained skills as indicators of (a) individual reading achievement or (b) successful programs. One danger is that excessive testing of constrained skills may lead to an

overemphasis on these skills to the exclusion of unconstrained skills such as vocabulary and comprehension. A second risk is that policymakers and the public may equate success on constrained skills with reading proficiency. This would create a minimum competency approach to reading assessment that does not adequately assess children's emerging use and control of literacy. ]

### **But isn't speed important in fluency?**

Allington (2009) echoing the work of Samuels (2007) reminds us that while each of the DIBELS subtests (as well as some other alternative tools) incorporates the word *fluency*, the tag simply is not valid. We always have to be careful of thinking that because something is labeled as one thing that it really is that thing. Opitz (2007) discusses that in any contemporary view of fluency, speed is only one element. Fluency also encompasses accuracy and prosody (expression). Newer views of fluency also remind us that fluency is not about reading as fast as possible but knowing when to read fast and when to slow down and that doing so can happen within a sentence or paragraph. What's really important about speed is helping students understand how to use it to best comprehend. Control becomes another key aspect of fluency. Finally, most views of fluency are not divorced from understanding. Speed without understanding is not really fluent reading.

While DIBELS attaches itself to the term fluency, its narrow view of fluency is simply completing tasks fast with less concern about accuracy, prosody, control and/or understanding, making it an inadequate measure of fluency. As Samuels (2007) concludes: "DIBELS tests...are not valid tests of the construct of fluency as it is widely understood and defined...The creators of



DIBELS are guilty of reification. By attaching the term *fluency* to their tests, they create the false assumption that that is what their tests measures.” In fact because fluency-based measures mainly assess the ability to complete tasks fast and sort students according to these rates, their correlations are primarily about speed of completing tasks and not the actual tasks themselves (for example, correlating saying nonsense words as an indicator of future successful reading performance.) Altwerger, Jordan, and Shelton (2007) affirm this conclusion pointing out that since the DIBELS subtests and the standardized measures with which they are often correlated are timed measures, it is time – not thinking or reading proficiency -- that is determining success or failure.

In addition, it’s not like we don’t know how to assess a more comprehensive view of fluency. NAEP includes a four point scale which can be used as a rubric for assessing fluency (<http://nces.ed.gov/nationsreportcard/studies/ors/scale.asp>). Opitz (2007) shows how to modify this scale for assessing fluency in the classroom and how these tools can be used to aggregate across learners to determine needs that can inform classroom instructional decisions (Opitz, Ford, & Ereksen, 2011).

**What does it matter if a tool is psychometrically flawed as long as it provides information about how students are doing?**

Remember that the most important goal of any RtI model is to make sure that all students have had access to quality instruction. Many tools do nothing to monitor that every child has been provided quality instruction. But what if we could assume that every student has had access to quality instruction? The next step is to monitor whether students are responding to instruction. Unfortunately, caution is needed when a single tool becomes the sole source of information used

to identify and sort students. When fluency-based measures assume this high stakes role, their flaws become more problematic. In addition to the problems that Pressley has surfaced, the major concerns come from the use of the results. Pearson (2006) cautions about the potential for “means-ends confusion.” In other words, a district must be careful that the tasks of fluency-based measures do not become ends in and of themselves. Curricula should not be circumvented so that getting better at the subtests end up replacing getting better at reading. High stakes decisions based on flawed assessment measures whether they are directed toward students or instruction have the potential of negatively impacting both. Pearson concluded: “I want [students] to make progress...because we provided them with rich curriculum that ensured balanced development of a range of skills and a broad exposure to important ideas – not because we had them practice time trials five times a week.”

**So how did one fluency--based measure become the RtI tool of choice for so many districts?**

Pearson is not alone in suggesting that this might have happened because of political positioning. As Pearson pointed out: “Of course, it does not hurt that DIBELS was officially blessed as a scientifically valid instrument...by the Reading First Assessment Academy, an advisory group on which DIBELS author Roland Good served.” One doesn’t have to take Pearson’s word for it. You can secure a copy of the Officer General Report on Reading First (2006) to see how certain decision-makers such as Roland Good, Ed Kame'enui, and Deborah Simmons involved in implementing Reading First policies had vested interests in favorably endorsing specific programs and products including DIBELS. Roland Good is the co-author of DIBELS and works at the University of Oregon. His colleagues at the time included Simmons and Kame'enui. Here’s how the report describes it: “All three of those former committee members - Roland

Good, Ed Kame'enui, and Deborah Simmons - benefited financially either directly or indirectly from the sale of a specific assessment product called the Dynamic Indicators of Basic Early Literacy Skills (DIBELS). Good was a co-author of DIBELS; so far, a company in which he owns a 50 percent share has received more than \$1.3 million in royalty and other payments from the sale of DIBELS.” This happened at a time when all three were involved in implementing Reading First policies. As we have told teachers repeatedly, “Follow the money!” For the past decade WSRA has spent a lot of time pointing out how blurry the line is between pedagogy, policy, politics and profits. This is just one more example of how blurry the line can get (Glenn, 2007; Brownstein & Hicks, 2006; Basken, 2006).

**So if we don't use DIBELS, how do we meet the demands for progress monitoring in RtI models?**

First, let's do everything possible to guarantee that all students are getting access to quality reading instruction. Assess the quality of instruction first! If we can be assured that all students have this access, than why not monitor progress by letting kids read and write? If one really wants to see whether the classroom reading instruction and/or school intervention programs are working, why not develop or select authentic reading and writing tasks that are compatible with your local standards, benchmarks, expectations, philosophy, materials and methods? This may be even more important as districts are influenced by the Common Core State Standards and held accountable by compatible assessments from consortia. Part of the attraction of some tools is that they provide the possibility of easy, cheap progress monitoring on a widespread scale, but it needs to be pointed out that this type of widespread progress monitoring is not required in RtI models. Districts can be much more creative and effective at developing monitoring systems that

accurately identify students who are not responding to instruction and intervention. Educators that want to look at alternatives may want to examine the article reprinted in this issue in which Linda Dorn and Barbara Schubert show how the principles at the heart of the Comprehensive Literacy Model can become a Comprehensive Intervention Model. In a previous *WSRA Journal* article, Salli Forbes (2008) also profiled educators from three districts across the country to show how they integrated existing Reading Recovery programs in their RtI plans.

### **What else do we need to know about RtI?**

When Susan Ohanian was asked about the public's need for accountability systems in education at the UW Reading Symposium, her response was something like this, "Don't ask me to be accountable for what goes on in schools until we can ask society to be accountable for guaranteeing that every child can live in a family that is working for a living wage." Like many other discussions in education, RtI is most often seen as a school-based response to problems that may transcend schools. In his recent article "Reading First and Its Impact on Poor Children" in *Dissent Magazine*, Gerald Coles said this: "Since the policy insists that everything needed to be done for poor children's education is being done, it also serves to justify doing as little as possible for poor children in every area of their lives that influences educational outcome." I was thinking about this after reading an article "Class Matters: In and Out of School" by Jane Boyd-Zaharis and Ghelen Pate-Bain (2008) in the September *Phi Delta Kappan*. In presenting a hierarchy of needs for a self-actualized society, the authors cause me to imagine a tiered-approach that first guaranteed the families of students affordable housing in stable neighborhoods, second a living wage with health care benefits, third access to quality early childhood education, and finally quality instruction in small classrooms. Would tiering

interventions related to those conditions that transcend schools also work to lower the identification and placement of students in special education? My best guess would be – that might even work better.

**Where can we get more information about looking at DIBELS, CBM and RtI more critically?**

For a comprehensive, critical look at RtI you may want to begin with **What Really Matters in Response to Intervention: Research-based Designs** by Richard Allington (Allyn & Bacon, 2009)

Here are some of the resources that have influenced the information shared in this article:

**Understanding and Using Reading Achievement K-12** by Peter Afflerbach (IRA, 2007)

“RTI Method Gets Boost In Spec. Ed.: Intervention Can Spot Learning Disabilities” by Christina A. Samuels in **Education Week** (November 30, 2005)

**RTI: Responses to Intervention: Fundamentals of Process & Practice** by Burggraf and Sotomayer (DayOne Publishing, 2007)

“Where We Are...Connecticut’s Response to Intervention (RtI) Model: Using Scientifically Research-based Interventions Improving Education for ALL Students Connecticut’s Framework for RtI” by Karen Costello in **WSRA Journal** (Summer 2008)

“Why Do Some Students Fail to Learn to read? Ockham’s Razor and the Conditions of Learning” by Brian Cambourne in **The Reading Teacher** (May 2001)

**Accessible Assessment: How 9 Sensible Techniques Can Power Data-Driven Reading Instruction** by Opitz, Ford and Ereksen (Heinemann, 2011)

“How Important is Reading Skill Fluency for Comprehension?” by Walczyk and Griffith-Ross in **The Reading Teacher** (March 2007)

**RTI: A Practitioner’s Guide to Implementing Response to Intervention** by Mellard and Johnson (Corwin, 2008)

**Reponses to Intervention: A Practical Guide for Every Teacher** by Bender and Shores (Corwin, 2007)

“Relationship between CCSS and RtI in Literacy and Language” by Karen Wixson & Marjorie Lipson in **The Reading Teacher** (March 2012)

**The Truth About DIBELS: What It Is - What It Does** edited by Kenneth S. Goodman

(Heinemann, 2006) [Note the foreword by P. David Pearson is one of the best summaries of the DIBELS critique.]

“National Cloud of DIBELS Test Draws Scrutiny: Critics Say Reading Tool’s Scope Fails to Justify Its Broad Use” by Kathleen Kennedy Manzo in **Education Week** (September 28, 2005)

“How Should Fluency Be Assessed?” in **What Really Matters in Fluency: Research-based Practices across the Curriculum** by Richard L. Allington (Allyn and Bacon, 2009)

**Evaluation of Reading First in Michigan Technical Report #1: Do Fluency Measures Predict Reading Achievement? Results from the 2002-2003 School Year in Michigan’s Reading First Schools** by Carlisle, Schilling, Scott and Zeng (Ann Arbor, MI: University of Michigan -- 2004) [Cited by Pressley in his critique of DIBELS]

“The Relation between DIBELS, Reading Comprehension and Vocabulary in Urban First Grade Students” by Brant Reidl in **Reading Research Quarterly** (October/November/December 2007)

“Reinterpreting the Development of Reading Skills” by Scott Paris reprinted in the **WSRA Journal** (Summer 2012).

“An Empirical Review of Psychometric Evidence for the Dynamic Indicators of Basic Early Literacy Skills” by Catherine T. Goffreda and James Clyde DiPerna in **School Psychology Review** (2010, 463–483).

**Don't Speed. Read: 12 Steps to Smart and Sensible Fluency Instruction** by Michael Opitz  
(Scholastic, 2007)

“The DIBELS Test: Is Speed of Barking at Print What We Mean By Reading Fluency, or  
Comprehension for that Matter?” by S J Samuels in **Reading Research Quarterly**  
(October/November/December 2007)

**Rereading Fluency: Process, Practice and Policy** by Altwerger, Jordan and Shelton  
(Heinemann, 2007)

Office of the Inspector General (2006). **The Reading First Program's Grant Application  
Process: Final Inspection Report** (No. ED-OIG/I13-F0017). Washington, DC: US Department  
of Education.

“Reading for Profit: Whistleblowers Allege that U of Oregon Scholars Steered Bounty from No  
Child Left Behind Act to Themselves and Their Colleagues” by David Glenn in **The Chronicle  
of Higher Education** (February 2, 2007)

“ED Ignored Early Warnings on Reading First Conflicts, Report Says: Officials Obscured  
Origins of Influential Assessment Review” and “Former Reading First Director Draws Fire –  
and Defenders: Doherty Says He Used ‘Bad Judgment But Upheld the Law’” by Brownstein and  
Hicks on **TitleIonline** (October 30, 2006)



“Lawmakers Asks Criminal Probe of US Reading Program (Update 1)” by Paul Basken on **Bloomberg.com** (September 25, 2006)

“A Comprehensive Intervention Model for Preventing Reading Failure: A response to Intervention Progress” by Linda Dorn and Barbara Schubert in **WSRA Journal** (Summer 2008)

“Reading Recovery: A Major Model of Many RTI Models” by Salli Forbes in **WSRA Journal** (Summer 2008)

“Reading First and Its Impact on Poor Children” by Gerald Coles in **Dissent Magazine** (September 4, 2008)

“Class Matters – In and Out of School” by Jayne Boyd-Zaharias and Helen Pate-Bain in **Phi Delta Kappan** (September 2008)