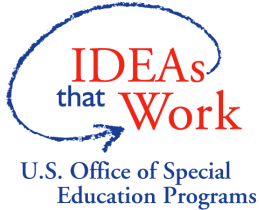




Responsiveness-to-Intervention Symposium

December 4-5, 2003 • Kansas City, Missouri

The National Research Center on Learning Disabilities, a collaborative project of staff at Vanderbilt University and the University of Kansas, sponsored this two-day symposium focusing on responsiveness-to-intervention (RTI) issues.



The symposium was made possible by the support of the U.S. Department of Education Office of Special Education Programs. Renee Bradley, Project Officer. Opinions expressed herein are those of the authors and do not necessarily represent the position of the U.S. Department of Education.

RTI: It's All About *The Nudge*

Donald L. Compton
Vanderbilt University

When citing materials presented during the symposium, please use the following: “Compton, D. L. (2003, December). *RTI: It's All About The Nudge*. Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.”

RTI: It's All About *The Nudge*

Jack Fletcher was recently awarded the Samuel T. Orton Award from the International Dyslexia Association (formerly the Orton Dyslexia Society). In his acceptance speech Jack talked about what I will refer to as “*the nudge*”. As Jack put it, “Reading is not natural and everyone is taught to read. Some of us require just a *nudge* from an adult and others require years of *nudges* and even that’s not enough.” The responsiveness to intervention (RTI) model of LD identification seems to be all about *the nudge*. Children who require a gentle nudge from an adult to learn to read are obviously not LD in the area of reading, whereas children requiring years of nudges to learn to read probably are LD.

So is it fair to equate *the nudge* with our current educational service delivery in which children receive a gentle nudge in general education and a much more substantial nudge in special education. Probably not, *the nudge* as it turns out is far more than just a synonym for our dichotomized educational delivery system. As an alternative to our current LD identification system, the RTI model requires us to delve much deeper into what is meant by *the nudge*. For instance, in order for a RTI model to become operational we must specify who should get *the nudge*, what *the nudge* should consist of, when *the nudge* should occur, how long *the nudge* should last, and by whom *the nudge* should be applied? And while this session’s title “How should unresponsiveness to secondary intervention be operationalized” seems innocent enough, we see in each of the three presentations an awareness by the author that unresponsiveness can not be separated from the who, what, when, for how long, and by whom of *the nudge*. It is clear from this sessions

presentations that for us to more fully understand what is meant by unresponsiveness in a RTI model of LD we are forced to consider it within the context of who, what, when, for how long, and by whom. I hope you will indulge me as I demonstrate how each of our presenters has grappled with these aspects of *the nudge* and how this shapes and perhaps challenges our present understanding of unresponsiveness.

Good, Simmons, and Kame’enui (2001) have recently suggested that “Assessment for educational prevention and accountability requires more than just a new test; it requires a different conceptual approach. In the primary grades, such an assessment system in schools at minimum must reliably (a) document and account for growth on a continuum of foundational reading skills, (b) predict success or failure on criterion measures of performance (i.e., high-stakes tests), and (c) provide an instructional goal that if met will prevent reading failure and promote reading success” (p. 261). Within the context of an RTI model of LD, Good et al. are proposing an assessment system that not only identifies who should receive a change in instruction (i.e., an additional *nudge*), but also when such instructional changes should be made and potentially how long such instructional changes should last. So we see that who should get *the nudge* is confounded by the when and the how long of *the nudge*.

Good and colleagues propose the use of a continuum of fluency-based indicators of foundational early literacy skills (Dynamic Indicators of Basic Early Literacy Skills, DIBELS) to predict reading outcomes, to inform educational decisions, and to change

reading outcomes for students at-risk of reading difficulties. To help operationalize the DIBELS assessment instrument and procedures a set of benchmarks have been established at each grade for the appropriate foundational early literacy skills. Performance benchmarks across the various early literacy skills are linked across grade levels with the intent of establishing a seamless set of performance indicators. These benchmarks are described as critical indicators of student performance that are specific, measurable, and ambitious. The intent of benchmark goals is to specify a level of performance where the odds of attaining subsequent goals are high while also maintaining a reasonable passing rate. The anchor for the system of benchmark goals is for all first-grade children to read at or above 40 words per minutes on grade-level material using CBM oral reading fluency (ORF) procedures at the end of the year. Good et al. argue “support for the benchmark goal of 40 words or more on CBM ORF in spring of first grade for all students derives from empirical, theoretical, and social-validation source” (266). Additionally, a score below 10 on CBM ORF in the spring of first grade appears to have utility as a level where intensive instructional support is needed. By intensive instructional support I assume that Good et al. are referring to an early intervention protocol similar to those developed and validated by the various researchers funded by NICHD, OSEP, and IES.

In sum, the concept of predictive benchmarks seems to make good sense as one means of operationalizing unresponsiveness. Good et al. (2001) have argued that a score of less than 10 on CBM ORF should serve as the cut-off for identifying children in need of intensive instructional support (i.e., unresponsive to the general education curriculum) and a range of scores from 10 to 39 on CBM ORF for identifying children as

at-risk, based on end of first grade performance. However, do these benchmark cut-off scores work uniformly well across schools in different districts, states, and regions? I doubt it. For instance, Good et al. indicated that 69% of kindergarten children enrolled in schools that were involved in a federally funded schoolwide early literacy program passed the kindergarten phonemic segmentation fluency benchmark verses only 21% in another research site not engaged in schoolwide reform. If we assume all else to be equal in this example, differences in kindergarten curriculum across the two sites resulted in 2.5 times more children being identified as potentially in need of *the nudge*. While this certainly speaks to the power of research-based early reading curriculums, it also graphically illustrates the potential for variation across sites when the same benchmarks are applied uniformly across varying sites. The question then becomes, what happens when the number of children identified as needing *the nudge* far outnumber those available to provide *the nudge*? Additionally, are benchmarks equally effective in making decisions as to when intensive instruction should be discontinued (i.e., the how long)? For instance, should the goal of intensive intervention for a child that scores less than 10 on the end of first grade CBM ORF be to reach the end of first grade benchmark of 40, and if so does this improve the child’s odds of future reading success to the same degree as children who achieved this goal with intensive intervention? Finally, are all interventions equally equipped to aid children in achieving the key benchmarks? I hope it is apparent by this first illustration that within an RTI model of LD identification it is virtually impossible to separate who should get identified as unresponsive from the what intervention was used, when the intervention occurred,

how long the intervention lasted, and who applied the intervention.

Vellutino and his colleagues take a somewhat different approach to identifying who should get *the nudge*. Vellutino administers a quick and reliable initial screener (letter-name knowledge) to children as they enter into kindergarten. Based on screening performance the lowest 30% of the sample is identified as at-risk for early reading difficulties. I do not believe that Vellutino and colleagues consider performance in the lowest 30% on letter identification as a benchmark in the way that Good and his colleagues would. In addition, Vellutino's selection criteria would vary across samples depending on the risk status of the school or district population (e.g., the amount of poverty, percentage of EL, etc.). The lowest 30% is selected because it helps to insure a low number of false negatives, and thus is probably an overestimate of the size of the actual risk pool. However, by limiting the number of false negatives it allows prevalence rates to be more accurately estimated.

This group of at-risk children is then enrolled in a structured intervention program. A certified teacher provides the group of at-risk children with 30 minutes twice a week of small group remediation focusing on emergent literacy skill. Remedial activities focus on emergent literacy skills such as knowledge of print concepts, print awareness, letter recognition, letter identification, phonological awareness, letter-sound mapping, sight word learning, shared and guided reading, and listening and reading for fun and enjoyment. Vellutino and colleagues then follow these children into first grade and provide additional daily one-to-one tutoring to those children that continued to show signs of unresponsiveness. At-risk readers in first grade were defined as children who were performing below the mid-point of the children who had received intervention in kindergarten. Thus, by

definition, 50% of the children in kindergarten intervention were designated as at-risk readers in first grade. This gated system allows children to qualify for additional intervention if necessary. In addition, we see the intensity of the intervention increasing over time as we concentrate the severity of the reading deficits in the at-risk group. Responsiveness at each grade level is then defined by performance in the normal range on a myriad of individually administered standardized tests. The promising news is that implementation of kindergarten intervention or kindergarten and first grade intervention combined can prevent reading difficulties in a majority of at-risk children at least through third grade.

At this point it is clear that in the Vellutino system *the nudge* evolves as the reading characteristics of the at-risk sample changes. In such a system it is safe to assume that the group of children judged to unresponsive by the end of first grade are those children in most need of added and more intense instruction. And so the Vellutino system of characterizing unresponsive children cannot be separated from the procedures, length of time, and intensity of the intervention. The logical question that then needs to be asked is, How much overlap is there between children identified by Good et al. using kindergarten and first grade benchmarks versus children identified in kindergarten and first grade using the gating system proposed by Vellutino and colleagues? My guess is that the children identified as unresponsive by the Vellutino system would also be identified as unresponsive by the Good benchmark procedure, but that the reciprocal is probably not true. The reason for this is that identification is intimately associated with intervention in the Vellutino system and somewhat less so in the Good system. It

is the attention to what intervention should consist of, when intervention should occur, how long intervention should last, and by whom intervention should be applied that makes the Vellutino system so appealing and probably so difficult to implement. We must somehow determine whether we lose the promise of the RTI identification procedure so elegantly demonstrated by Vellutino et al. by simplifying it to a set of highly reliable and easy to administer benchmarks as amply developed by Good and colleagues.

Finally, Torgesen pushes the envelope of our current understanding of RTI methods by asking how these models for LD identification can be used with older students. What becomes clear from Torgesen's discussion is that the who, what, when, and how long of *the nudge* changes as children grow older. Assuming that successful implementation of a RTI model in the primary grades will identify the majority of children who have phonological- or experience-based reading problems. The type of children most likely to be identified as candidates for special education in the upper grades are children who make inadequate progress in the area of reading comprehension. Scarborough (2003) has recently published data indicating that those identified in the upper elementary grades with reading problems have adequate decoding skills with deficits in the area of reading comprehension. Operationalization of an RTI model of LD identification to growth in reading comprehension forces us into uncharted territory. There is general agreement that the assessment of reading comprehension skills in developing readers is poorly understood. As an example of just how high the measurement hurdle is, I am in the process of conducting an intervention study intended to improve the reading comprehension skills of poor readers. At pretest we administered the 52 children (grades 3 to 5) who had been nominated by their teachers as poor in reading

comprehension the GORT-4 reading comprehension test and the WJRM-R: Passage Comprehension test. The correlation between standard score performance across the two measures was $r=.20$. Such a low correlation is enough to question whether in this sample these two well-known measures of reading comprehension are measuring the same construct. Making matters worse is the expectation that the existing measures of reading comprehension will be sensitive to growth in reading comprehension as a result of some kind of intervention seems to be pushing our present comprehension measurement system well past their capability. Without accurate measures of responsiveness to instruction the RTI model applied to older children with reading comprehension deficits is sure to fail.

Assuming the psychometric challenges of accurately measuring change in reading comprehension skill can be overcome, we are still without validated treatment protocols for improving the language and comprehension skills necessary to impact reading comprehension skills. Torgesen has shown that the gains in comprehension skill (ranging across studies from .07 to .32 standard scores per hour of remediation) for older elementary children with LD enrolled in intensive intervention lag far behind than gains made by the same children in phonemic decoding skills (ranging across studies from .18 to .47 standard scores per hour of remediation) We will also be forced to make the difficult decision of whether low verbal ability, often the source of poor reading comprehension skill, constitutes a defensible category under the LD label. Torgesen's paper seems to suggest that there are significant issues still waiting to be addressed before a RTI model can be operationalized for the identification of older children with LD. At this point *the*

nudge for older children can at best be described as a *gentle tap*.

I hope that you have found the three presentations that make up this session as interesting and thought provoking as I have. I also appreciate your indulgence in my playful use of *the nudge*. I also hope you have a greater appreciation for how intimately related the definition of unresponsiveness and intervention are within a RTI model of LD identification. To define one without the other is the equivalent of asking what chances are of Vanderbilt winning its next basketball game, this analogy is for our hosts from the University of Kansas, without knowing whether we are talking about the men's or women's team. Thanks you.