Early Screening Is at the Heart of Prevention

Early intervention works. Because it is also expensive, it's important to be able to identify the kids who are most at risk of reading failure. Thanks to a new generation of screening assessments, we can identify these students as early as kindergarten—and then invest in interventions for them. The new assessments are brief, trustworthy, and easy to administer. They can be administered to all kindergartners through third-graders a few times a year, allowing teachers to identify which students need extra help. They take only five to ten minutes per child to administer and can typically be given by classroom, reading, or special education teachers or aides. Once identified, these students can receive the assistance they need, and the downward spiral that results from weak early reading skills can be averted.

Selecting Assessments for Your School

Today, schools actually have more than two dozen individually administered early screening assessments to choose from that are appropriate for kindergarten through third grade. These assessments cover a variety of reading skills—phonemic awareness, phonics, fluency, vocabulary, and reading comprehension—as well as a variety of assessment purposes: screening, diagnosis, and progress monitoring. (There are also outcomes assessments that educators should be aware of if they plan to conduct studies of the effectiveness of their reading programs.) As an introduction to the kind of information that teachers can glean from early reading assessments, examples from TPRI and Dynamic Indicators of Basic Early Literacy Skills (DIBELS) are included here. The figure (immediate right) shows some items from the TPRI Screening Section for children at the beginning of the first grade. The figure on page 15 shows how a kindergarten student's progress is monitored using DIBELS.

Fortunately, a team of researchers has reviewed the current crop of assessments, identified which have sufficient reliability and validity, and developed a Web site for educators that clearly indicates which assessments are appropriate for differing grades, skills, and purposes (available at **http://idea.uoregon.edu/assessment/**).

Most currently available assessments identify children using national norms. So, for example, schools can decide to intervene with all children who score in the bottom 10 to 20 percent nationally. Of course, just how many students this will be varies greatly by school. A few assessments have established benchmarks, or cut scores, that represent evidence-based thresholds indicating the likelihood of reading success (or failure), and recommend that schools intervene with all students who fall below the benchmark.

A key issue that has arisen during the instrument-development research is creating accurate instruments that are not too long. Assessment developers have been grappling with the fact that longer assessments provide more detailed data, but shorter assessments are more practical for the classroom. This led to differentiating between screening and diagnostic assessments. Typically, screening instruments tend to be short, taking as little as five to 10 minutes per child, and they identify which students are at risk or behind, as well as some information on which skills the

*Shown above are assessments and benchmarks from the Screening Section of the TPRI, the first early reading assessment to be used throughout a state. These two Screenings are used at the beginning of first grade along with a 10-item screening of children's knowledge of letter names and sounds. As you can see, the TPRI provides empirically derived criteria to indicate if students have developed adequate knowledge and skills. When students do not meet those criteria, the teacher moves directly into more in-depth assessments from the Inventory Section of the TPRI. Therefore, the amount of assessment is individualized: Students who are "developed" on the Screenings will be done in just five minutes; students lacking skills will continue through the Inventory assessments until the skills that need to be developed are identified—a process that can take an experienced teacher anywhere from 10 to 25 minutes. (For more information on the TPRI, see **www.tpri.org**.)*

How Do They Work?

The key to our new ability to predict which children are likely to have problems in learning to read is the research finding that almost all struggling readers have problems with phonemic awareness—identifying and being able to manipulate the sounds in words (Torgesen, 1998). Not surprisingly, given their troubles with the phonological features of language, these children also have difficulty grasping the alphabetic principle and are slow to build up a "sight vocabulary," meaning words that they can read automatically without sounding them out. Building on these highly consistent findings, researchers have found that by midway through kindergarten (assuming prereading skills are being taught), knowledge of letter names predicts future

students are lacking. Diagnostic instruments—used only for the smaller group of students deemed at risk in the screening—tend to be longer, taking roughly 20 to 45 minutes per child, and they offer a much more thorough look at students' strengths and weaknesses. (These time estimates are for teachers who are experienced in using these assessments. More time will be needed while teachers become accustomed to using those tools.) Sometimes assessments have both screening and diagnostic components. For example, the TPRI has "Screening" and "Inventory" sections. When children don't meet criteria in the Screening Section, the teacher can immediately switch to a more in-depth assessment from the Inventory Section to pinpoint the knowledge and skills that the child still needs to develop. Teachers can also use Inventory data to match instruction with specific student needs.

A third type of assessment is for progress monitoring. These instruments typically come in short, multiple forms so that students' skills can be assessed every two weeks (or even more frequently) to quickly determine if an intervention is sufficiently effective. If not, the intervention can be altered (by changing the instructional content, methods, and/or intensity), the child may be given a diagnostic assessment, or the child may be referred for special education. For example, DIBELS is a widely used screening and progress-monitoring assessment. DIBELS measures take just a few minutes each and usually come in 20 alternate forms for frequent checkups. The figure (left) shows how a kindergarten student's progress is monitored using DIBELS.

While screening instruments are used with all students, diagnostic instruments are only necessary for students whose screenings reveal serious skill deficits and/or whose progress monitoring indicates that they are not responding to the intervention. Ideally, all K-3 students should be screened three times per year starting in mid-kindergarten; diagnostic and progress monitoring assessments can be done as needed, with progress monitoring of children in an intervention being quite frequent to make sure that interventions are as effective as possible.
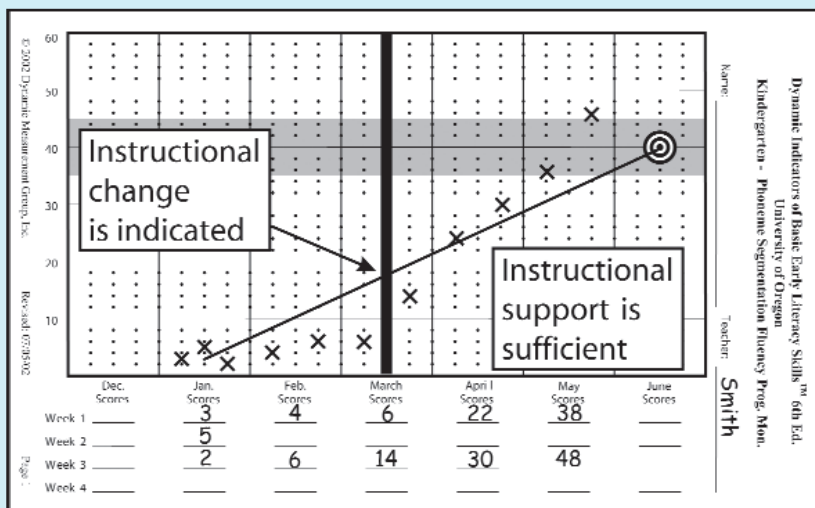
This chart shows one kindergarten student's improvement based on a progress monitoring assessment from January to June. Using alternate forms of the DIBELS measure of Phoneme Segmentation Fluency (PSF), the teacher screened this child three times in January to be sure that he really needed an intervention. As noted by the horizontal line in the chart, all students should score at or above 40 on PSF by the end of kindergarten. Drawing a line from the student's initial scores to that benchmark creates clear goals for the rest of the year and allows the teacher to judge the success of his intervention. Initially, the intervention was not sufficient. The teacher made his intervention more intense by providing additional modeling, examples, and practice—and the student reached the PSF benchmark by the end of May. (To learn more about DIBELS, visit **http://dibels.uoregon.edu/** *and* **www.dibelsassessment.com**.*)*

reading ability. And by first grade, letter-sound knowledge is highly predictive. (For more on this topic, see "Catch Them Before They Fall: Identification and Assessment To Prevent Reading Failure in Young Children" from the Spring/Summer 1998 issue of *American Educator;* it is available online at **www.aft.org/pubs-reports/ american_educator/spring_sum98/ torgesen.pdf**.)

How Accurate Are They?

Just how accurate are these early assessments? Accuracy varies by instrument. Rather than reviewing several assessments, let's look at the average predictive power of assessing kindergartners' letter identification skills (Snow et al., 1998). A meta-analysis of 20 studies that measured 11 different possible predictors of reading difficulties (including receptive vocabulary, expressive language, concepts of print, and verbal memory of stories or sentences) found that letter identification was the strongest single indicator of future reading. The mean correlation between letter identification in kindergarten and reading scores in grades one through three was .52. In fact, letter identification was almost as good a predictor by itself as an entire reading-readiness test (which includes a whole host of reading skills). But what does a moderately strong correlation like this mean when it comes to designating children at risk or not? Another study (Snow et al., 1998) used 1,000 kindergartners' letter identification skills to find out. The researchers considered their predictions accurate if the children who were designated at risk in kindergarten were then in the bottom 20 percent on teachers' ratings in first grade.

To begin with, the researchers tested a strict letter-identification cutoff; they designated students at risk only if they fell in the bottom 10 percent. According to the first-grade teachers' ratings, this strict cutoff correctly identified 83.2 percent of children. Since there were 1,000 children in the study and the bottom 10 percent were designated at risk, 100 children were so designated. Of these, 63 were correctly identified (meaning they were in the bottom 20 percent according to teachers' ratings in first grade), but 37 were false alarms (meaning they were not in the bottom 20 percent). Of the 900 children designated not at risk, 769 were correctly identified, but 131 were

misidentified (meaning they were in the bottom 20 percent in first grade).

Believing that too many children who did end up having reading difficulties were missed with the strict cutoff, the researchers also examined a more lenient letter-identification cutoff. In this second analysis, they designated the bottom 25 percent of kindergartners at risk. Of these 250 children, 118 were correctly identified, but 132 were false alarms. Of the 750 children designated not at risk, 677 were correctly identified, but 73 were not. Overall, the more lenient cutoff meant that the overall accuracy of the prediction was reduced slightly (79.5 percent of children were correctly identified)—but the percentage of struggling readers who were missed dropped from 15 to 11.

Obviously, educators have to make a conscious choice when they decide what percentage of children to intervene with. Intervening with the bottom 10 percent means that many at-risk children will not be appropriately served. And intervening with the bottom 25 percent means that many not-at-risk children will be served.

No assessment can completely overcome these potential errors in identifying at-risk children. Even with the best assessment, some children who will have reading problems are not identified and some who will not are. But there are strategies to greatly reduce the errors in identification. To minimize under-identification, schools are encouraged to screen all children—three times per year—starting with mid-K. (Assessments at the very beginning of kindergarten tend to be unreliable because students may lack skills simply because they haven't been taught, not because they will have trouble with the concepts once they have been presented in the regular classroom setting.) To minimize over-identification, assessments often come with multiple forms so that teachers can confirm the results (and be sure that the child was not just having a bad day) before the intervention begins. Given the importance of addressing skill deficits, over-identification of children may be the best policy. For not-at-risk students, the intervention will simply reinforce their skills, acting like an "insurance policy" against future problems with reading. And, with adequate progress monitoring, such students will test out of the intervention quickly.

Fortunately, predictions of which students are at risk for reading failure become even more accurate by the end of first grade. This is what one would expect given that, starting at the end of first grade, students' word-reading ability can be assessed directly instead of indirectly through such pre-reading skills as letter naming and phoneme segmentation. While it is clearly true that early word reading ability is a strong predictor of later word reading ability, very brief measures of oral reading fluency are also a strong predictor, and thus a good screening measure, for difficulties in reading comprehension. In fact, Fuchs, Fuchs, Hosp, and Jenkins (2001) reported evidence that a very brief measure of oral reading fluency was a better predictor of performance on a reading comprehension outcome measure than was a brief measure of reading comprehension itself. In this study, with middle and junior high school students with reading disabilities, the correlation between oral reading fluency and the reading comprehension measure was a nearly perfect .91.

More recently, researchers comparing third graders' performance on the Dynamic Indicators of Basic Early Literacy Skills measure of Oral Reading Fluency to their scores on state assessments of reading comprehension have found correlations of .70 with the Florida Comprehensive Assessment Test (Buck and Torgesen, 2003) and .73 with the North Carolina end-of-grade assessment (Barger, 2003).

—*EDITORS*

References

Barger, J. (2003). Comparing the DIBELS Oral Reading Fluency indicator and the North Carolina end-of-grade reading assessment. Asheville, N.C.: North Carolina Teacher Academy.

Buck, J. and Torgesen, J. (2003). The Relationship Between Performance on a Measure of Oral Reading Fluency and Performance on the Florida Comprehensive Assessment Test. Tallahassee, Fla.: Florida Center for Reading Research.

Fuchs, L.S., Fuchs, D., Hosp, M.K., and Jenkins, J.R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5,* 239-256.

Snow, C. E., Burns, M. S., and Griffin, P., Eds. (1998). "Predictors of Success and Failure in Reading" in *Preventing Reading Difficulties in Young Children.* Washington, D.C.: National Academy Press.

Torgesen, J. K. (1998). "Catch Them Before They Fall: Identification and Assessment to Prevent Reading Failure in Young Children, *American Educator, 22 (1&2),* 32-39.