Using Oral Reading Fluency as a Benchmark Assessment

by Jan Hasbrouck

Curriculum-based measurement (CBM) has a research base that extends over 25 years (see "Progress Monitoring Within a Multitiered Prevention System" by Lynn S. Fuchs and Doug Fuchs in this issue) to support its use by teachers to identify students who may need extra help with basic academic skills and to monitor their progress as they receive instruction. CBM for reading, or CBM-R, is a valuable set of procedures that can be used by teachers and specialists to make decisions about students' progress—or lack of progress—in reading.

The act of reading involves processing phonological (phoneme awareness), orthographic (decoding skills), semantic (vocabulary), and context (background knowledge) information through the application of decoding, word recognition, and fluency skills to understand the

meaning of written text. The research on CBM-R indicates that results from these quick, fluency-based measures, when administered using a set of standardized procedures, can be interpreted as an accurate indicator of overall skill level in the complex process of reading, not just as an indicator of the development of the fluency component of reading.

CBM-R involves assessing students' oral reading fluency by measuring *words read correctly per minute* or *WCPM* (see "Using CBM-Reading Assessments to Monitor Reading Progress" in this issue by Joseph R. Jenkins and colleagues). Educators use these procedures in two ways to monitor their students' progress. One procedure is to assess students three times per year and compare the resulting scores to a benchmark score. This procedure, commonly called screening or benchmarking, ensures that all students

make sufficient academic progress over the course of the school year to stay on-track for future skill development. If a student's score does not reach the expected benchmark at any point, teachers know to take a careful look to determine if that student might benefit from extra instructional assistance.

A second form of CBM-R is typically used only with students who are already receiving extra assistance with their reading at a supplementary (Tier 2) or intensive (Tier 3) level that would include many students with dyslexia. (See "Progress Monitoring Within a Multitiered Prevention System" in this issue for a more detailed description of the tiers in a multitiered system.) This form of progress monitoring involves using CBM-R procedures as frequently as once a week. Teachers and specialists graph students' oral read-

ing fluency scores and then use those graphs to evaluate the effectiveness of each student's instructional program. (See "Using CBM-Reading Assessments to Monitor Reading Progress" in this issue for more information on this use of CBM-R.)

Defining Benchmarking

Benchmarking is widely used in elementary schools across the country to find students who need extra assistance with reading. Student progress is monitored from fall to winter to spring by having individual students read aloud one, two, or three unpracticed passages at their grade level for 1 minute to a teacher or other trained examiner. The number of passages varies depending on the assessment instrument. Students in kindergarten and the first half of first grade read aloud from lists of letter names and letter sounds for 1 minute. Other benchmark-

ing measures have students identify phonemic elements of words presented verbally.

Most schools using benchmarking procedures administer the assessments to every student, despite skill level. This approach ensures that students do not "fall between the cracks" and miss the assistance they need with their reading. This is especially important for students with dyslexia because current research strongly suggests that early intervention can prevent most reading difficulties.

How to Administer Benchmarking Assessments

During the administration of benchmarking assessments, the examiner listens for errors. Each word that is mispronounced, substituted for another word, or omitted is counted as one error. If a student hesitates on a word for more than 3 seconds, that is considered an error. Errors that are self-corrected, words that are read correctly but repeated, and words that are inserted by the student are not counted as errors. A passage score is determined by subtracting the total number of errors from the total number of words read in 1 minute. The resulting score is reported as WCPM. If two or three passages were administered, a final score would be calculated as either the average (mean) or the middle score (median) of the scores. This final score is then compared to a benchmark.

How to Interpret Benchmarking Scores

Some commercially available benchmarking tools such as *DIBELS* (Dynamic Indicators of Basic Early Literacy Skills) and *AIMSweb* have established their own recommended benchmarks by compiling scores from students. Other benchmarking tools, such as *The Reading Fluency Benchmark Assessor,* use the oral reading fluency norms compiled by Jan Hasbrouck and Gerald Tindal (Hasbrouck & Tindal, 2006) (see Figure 1).

> *...current research strongly suggests that early intervention can prevent most reading difficulties.*

Benchmark scores are calculated by comparing the performance of a group of students on a benchmarking assessment to their future performance on a comprehensive standardized assessment. A benchmark is determined to be the score that predicts which students will likely pass these comprehensive tests. For example, the benchmark score for the DIBELS assessment in the spring of first grade is 40 WCPM. This benchmark means that a student reading at least 40 WCPM on the DIBELS first grade passage in the spring will likely go on to do well with learning to read and to pass reading assessments in the future. If a student does not reach that benchmark of 40 WCPM, that student will probably need some extra assistance with reading. Additional assessments may also be necessary to determine why this student is falling behind.

The Hasbrouck and Tindal (2006) oral reading fluency norms can be used to make similar decisions. These norms were compiled from the WCPM scores of over 255,000 students from across the country, each of whom was given a benchmarking assessment using the standardized CBM-R procedures (unpracticed, grade-level passages read aloud for 1 minute). These norms simply describe the oral reading performance of a very large number of students who read a wide variety of passages. Hasbrouck and Tindal recommend using the 50th percentiles on their chart for benchmarking decisions. If a student's WCPM score is within 10 +/- of the 50th percentile score in the fall, winter, or spring, a teacher can assume that that student is likely on track with reading. A WCPM score more than 10 words below the 50th percentile may indicate that the student is having difficulty with his or her reading and further actions may need to be taken. To be extra vigilant about these important decisions, Hasbrouck and Tindal recommend that WCPM scores that fall in the range between 5 and 10 words below the 50th percentile be considered a "yellow flag" for a student. A student with a WCPM in this range may be on the way to having difficulty with reading and should be observed carefully during instruction.

GRADE	PERCENTILE	FALL WCPM	WINTER WCPM	SPRING WCPM
1	90		81	111
	75		47	82
	50		23	53
	25		12	28
	10		6	15
2	90	106	125	142
	75	79	100	117
	50	51	72	89
	25	25	42	61
	10	11	18	31
3	90	128	146	162
	75	99	120	137
	50	71	92	107
	25	44	62	78
	10	21	36	48
4	90	145	166	180
	75	119	139	152
	50	94	112	123
	25	68	87	98
	10	45	61	72
5	90	166	182	194
	75	139	156	168
	50	110	127	139
	25	85	99	109
	10	61	74	83
6	90	177	195	204
	75	153	167	177
	50	127	140	150
	25	98	111	122
	10	68	82	93
	90	177	195	204
7	90	180	192	202
	75	156	165	177
	50	128	136	150
	25	102	109	123
	10	79	88	98
8	90	185	199	199
	75	161	173	177
	50	133	146	151
	25	106	115	124
	10	77	84	97

Figure 1. WCPM = Word Count Per Minute

The following case provides a useful example of the screening decision process. Jessica is in second grade. Her teacher administered a fluency-based benchmarking assessment to each of her students in February to determine if they were making sufficient progress in reading at the midpoint in the school year. She followed standardized procedures and used three different second-grade-level passages that had not been previously read by any of her students for this screening. Jessica's scores on the three assessments were 63, 57, and 59 WCPM. The teacher compared Jessica's median score of 59 WCPM to the Hasbrouck and Tindal (2006) second grade winter scores. The score of 59 WCPM falls 13 words below the 50th percentile score of 72 WCPM, but is 17 words above the 25th percentile score of 42 WCPM. This result indicates that Jessica may not be making adequate

progress in reading. Further assessment is warranted to verify this conclusion, in addition to examining Jessica's performance in daily reading lessons and independent work. The teacher realized that Jessica was showing increased reluctance to read, and when called upon to read aloud or answer questions, she often made errors, read with hesitation, or seemed confused by the text. Jessica's teacher decided to investigate further, by conducting one-to-one, second-grade-level skills assessments with Jessica.

Concerns About Benchmarking

Some educators have expressed very strong concerns about benchmarking assessments. How can such a short—60 second—measure, of only one, isolated reading skill—fluency—be used to determine a student's progress in the highly complex linguistic act of reading? (Hamilton & Shinn, 2003). This concern is certainly legitimate and can be addressed in two ways: first, by considering logically the purpose of benchmarking, and second, by examining the research base supporting these assessments.

Purpose of Benchmarking. Educators use oral reading fluency in much the same way that physicians use thermometers. Both are measures that provide a quick "score" that has scientifically proven reliability and validity; that is, when either benchmarking assessments are used to measure student performance (WCPM), or a thermometer is used to measure body temperature (Fahrenheit or Celsius), both provide "scores" that are consistent, accurate, and useful. Moreover, in both cases, the score obtained is compared to a calculated benchmark that is then used as an indicator of general "wellness" (on-track for reading) or "illness" (may need some extra assistance with reading).

Even when using such a precise tool, physicians understand that body temperature does not tell the whole story. If a man comes to an emergency room with a serious injury to his leg, it is likely that one of the first things that will happen is that someone will take his temperature. If it turns out to be in the normal range, the man will not be sent home, because the medical staff is trained to treat body tem-

perature as only one, single indicator of health or wellness. On the other hand, if this same man had a body temperature of 103.7 degrees Fahrenheit, a doctor would not rush him off to surgery to remove his gall bladder. Along with the leg injury, this person may have something else happening in his body. A thermometer reading does not say what is wrong, only that something is wrong. It is only an indicator that the physician should look carefully for a variety of possible causes. In much the same way, it is important for educators to understand that benchmarking assessments provide a score that can be used to indicate academic progress, but professional judgment is necessary for interpreting the score and deciding how to respond.

Research Base. Benchmarking is supported by 25 years of research. The article "Progress Monitoring Within a Multitiered Prevention System" by Lynn and Doug Fuchs, in this issue, provides an overview of the research base in CBM that provides strong scientific support for using benchmark assessments for making decisions about students' progress in reading. Educators who are concerned about using these assessments should review this research to understand why benchmarking assessments were developed and how to interpret them.

Summary

Oral reading fluency measures have a well-established role in the assessment of students for various purposes. One way to use these measures is for screening a large number of students to determine which ones have achieved a "benchmark" or a certain level of performance, and which students have not yet achieved that level of skill development. When the students who are lagging behind have been identified, appropriate interventions can be determined as quickly as possible. It is vital that as professional educators—teachers and administrators—we understand the tools available for making key decisions about our students. Quick, efficient, and accurate benchmarking assessments can help us find potentially struggling readers as quickly as possible and respond with appropriate instruction. Benchmarking is certainly one important tool that should be in the professional repertoire of every teacher who works with students with dyslexia and other reading problems.

Resources for CBM-R, Benchmarking, and Progress Monitoring

DIBELS (Dynamic Indicators of Basic Early Literacy Skills), http://dibels.uoregon.edu

EasyCBM System, http://easycbm.com

Edformation (2004). *AIMSweb* progress monitoring and assessment system, http://www.edformation.com/

National Center on Student Progress Monitoring, http://www.studentprogress.org

Read Naturally (2002). Reading Fluency Benchmark Assessor and Reading Fluency Progress Monitor, http://www.readnaturally.com

Shinn, M. R. (Ed.). (1989). Curriculum-based measurement: Assessing special children. NY: Guilford.

Shinn, M. R. (Ed.). (1998). Advanced applications of curriculum-based measurement. NY: Guilford.

References

Hamilton, C., & Shinn, M. R. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review, 32*(2), 228–240.

Hasbrouck, J. (2006). Putting fluency in perspective. *Balanced Reading Instruction, 13*, 9–22.

Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher, 59*(7), 636–644.

Jan Hasbrouck, Ph.D., *is president of JH Consulting. She provides training workshops and consultation across the United States and internationally on strategies for struggling readers and reading coaches. Dr. Hasbrouck worked as a reading specialist for 15 years and taught undergraduate and graduate courses as a university professor at Texas A&M University and the University of Oregon. Her research in reading fluency, academic assessment and interventions, and coaching and consultation has been widely published. Dr. Hasbrouck c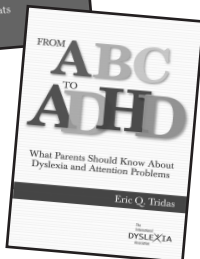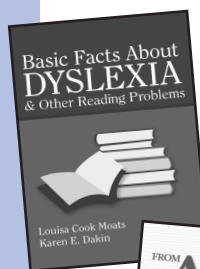an be contacted through her website www.jhasbrouck.com*