Ohio Technical Report

The Relationship Between Oral Reading Fluency and Ohio Proficiency Testing in Reading

Carolyn D. Vander Meer,

Fairfield City School District

F. Edward Lentz,

University of Cincinnati

Stephanie Stollar

Southwest Ohio Special Education Regional Resource Center

Suggested Citation:

Vander Meer, C. D., Lentz, F. E., & Stollar, S. (2005). *The relationship between oral reading fluency and Ohio proficiency testing in reading* (Technical Report). Eugene, OR: University of Oregon.

The Relationship Between Oral Reading Fluency and Ohio Proficiency Testing in Reading

Introduction

The standards-based reform movement and research on preventing reading failure have come together to present a terrific opportunity to improve academic outcomes for all children. In Ohio, the English Language Arts Content Standards guide curricula and instruction and provide the foundation for the state's assessment and accountability system. Unlike many states, at the time these data were collected, reading measures were administered multiple times during fourth grade. With the passage of No Child Left Behind, Ohio is moving from a fourth grade proficiency test to a series of achievement tests, the first of which was on reading in third grade. In addition to the state's accountability system, many Ohio schools use the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) to identify students for intervention and to monitor the effectiveness of intervention.

Purpose of the Study

The purposes of the current study were to a) examine the end of third grade and beginning and end of fourth grade oral reading fluency goals set forth by Good and Kaminski (2002) in comparison to Ohio expectations for fourth grade reading proficiency, and b) explore correlations between oral reading fluency and the reading portion of the Ohio Proficiency Test (OPT). These DIBELS benchmark goals can provide important targets for instruction in third and fourth grade. This study sought to confirm the connection between achieving the DIBELS benchmark fluency goals and passing Ohio's Fourth Grade Reading Proficiency Test.

Subjects

The DIBELS and OPT data were collected from three elementary schools in southwest Ohio. The schools were three of five elementary schools in a suburban school district of approximately 8,800 students. The schools house first through fourth grade students. These schools were chosen because of the availability of data for grades 3 and 4.

A total of 364 students who were in third grade during 2001-02 and fourth in 2002-03 participated in the study. All students with the exception of those identified with significant cognitive disabilities were included in the study. Students with an Individualized Education Program (IEP) were provided allowable accommodations during the test. The sample size varies across different analyses due to students moving in or out of school or due to absences. Table 1 summarizes the demographic information for each school involved in the study.

Table 1

Demographic Information of Participating Schools

		Percentage						
		Ethnicity					School	
School	Total Enrolled	Asian	Black	White	Hispanic	Multi-racial	Economically Disadvantaged	Students with Disabilities
A	553	4.0	15.4	69.5	6.1	4.3	22.2	9.5
B	556	2.0	2.9	93.8	0.4	0.9	8.3	7.8
C	508	1.8	7.4	85.6	1.0	4.3	12.4	10.4

Measures

Three measures of student academic performance were used in the present study: (1) Curriculum-Based Measurement Oral Reading Fluency (CBM ORF), (2) DIBELS Oral Reading

Fluency (DORF), and (3) Ohio Fourth Grade Reading Proficiency Test (OPT). DIBELS and CBM ORF data were reported for grades 3 and 4, respectively. OPT data was reported for grade 4.

Curriculum Based Measurement Oral Reading Fluency (CBM ORF): CBM ORF is a standardized, individually administered test of accuracy and fluency with connected text. The assessment consists of three reading passages, each of which students read aloud for one minute. Errors are words omitted, substitutions, and hesitations of more than three seconds. Words self-corrected within three seconds are not marked as an error. The oral reading fluency rate is the number of correct words per minute from the median of three passages (Good, Simmons, & Kame'enui, 2001).

The fourth grade students were administered CBM ORF probes in the fall and spring which were developed from randomly selected passages taken from the Houghton Mifflin Reading Series (Durr & Pikulski, 1986). All reading probes were administered individually by trained staff following the standardized procedures described by Shinn (1989).

A number of studies on oral reading fluency have confirmed the technical adequacy of CBM ORF procedures (Good & Jefferson, 1998; Tindal, Marston, & Deno, 1983). Large overviews of the research have validated CBM ORF measures as an excellent overall measure of reading achievement (Marston, 1989; Shinn, 1989; 1998). Further information regarding the reliability and validity of ORF can be found at http://dibels.uoregon.edu.

Measures of oral reading fluency from curriculum-based passages are used as indicators of reading achievement (Shinn, 1997). Students' performance on these indicators is then compared to performance expectation, or "where we would expect children to perform," to identify children at risk of reading disability. Performance expectations may be derived from

two sources: local normative data or performance associated with early reading success (Kaminski and Good, 1996).

DIBELS Oral Reading Fluency (DORF): DORF measures are also standardized individually administered measures of accuracy and fluency with connected text. The probes were created for the DIBELS measurement system and controlled for readability. The procedures for administration and scoring were the same as those described above for CBM ORF. All reading probes were administered individually by trained staff following the standardized procedures described above (Good, Kaminiski, and Dill, 2002; Shinn, 1989). The third grade students participating in this study were administered DORF probes in the fall, winter, and spring.

Benchmark goals for achievement of reading and indicators of risk for kindergarten through third grade have been derived from previous research (Good et al., 2003; Good, Simmons, & Kame'enui, 2001). Estimated grade 4 goals and risk indicators were based on CBM normative information from fourth grade students in fall, winter, and spring from Hasbrouck & Tindal (1992) and the slope of reading progress information from Fuchs, Fuchs, Hamlett, Walz, & Germann (1993). Further information regarding the development of the goals and indicators of risk can be found at http://dibels.uoregon.edu. DIBELS benchmarks for kindergarten through third grade and fourth grade CBM normative data were used by the schools in this study to identify at-risk students who were in need of intervention, to monitor progress, and for program evaluation. Table 2 provides the benchmark goals and risk indicators used in this study for grades 3 and 4 that were taken from the DIBELS website.

Table 2

ORF Grade 3 and Grade 4 Benchmark Goals and Indicators of Risk.

Indicator	Score		
	Grade 3 Spring	Grade 4 Fall	Grade 4 Spring
Low risk	110	93	118
Some risk	80 – 109	71 – 92	96 – 117
At risk	0 – 79	0 – 70	0 – 95

Ohio Fourth Grade Reading Proficiency Test (OPT): The portion of the OPT used in the present study is the Fourth Grade Reading Proficiency Test. The purpose of this test is to certify a fourth grade level of literacy and to identify students who are in need of intervention. The test is defined by twenty learning outcomes. Committees made up of Ohio citizens (educators and business members) through a consensus-building process developed these learning outcomes, based on Ohio's Model Competency-Based Language Arts Program.

The learning outcomes define the proficiencies fourth grade students are expected to possess and apply as a result of their learning experiences in kindergarten through February of the fourth grade year. The twenty outcomes are grouped into four subscales: 1) Constructing Meaning: Fiction, 2) Examining/Extending Meaning: Fiction, 3) Constructing Meaning: Nonfiction, and 4) Examining/Extending Meaning: Nonfiction. The subscale performance of students was used to identify specific student needs for decisions related to intervention and curriculum.

The items on the reading test are based on fiction, poetry and nonfiction reading selections. A total of 30 items are counted to obtain the students' reading scores. Each form of the proficiency tests includes multiple-choice items, short-answer items, and extended-response items. The number of each type of item is determined by the reading selections in the form.

Multiple-choice questions included on the Fourth Grade Reading Proficiency Test emphasize critical thinking rather than factual recall. Each reading selection has at least one short-answer or one extended-response item. Further proficiency test development and statistical information is located on the Ohio Department of Education website at http://www.ode.state.oh.us/proficiency/.

The Fourth Grade Reading Proficiency Test (OPT) was administered statewide to all students in grade 4. Students were given the opportunity to take this test three times during the school year. The first test administration was given in October 2002 and was mandatory for all students. Two additional test opportunities were given in March and July 2003. This school district elected to administer the test in March to those students who scored within the proficient range or below on the October OPT. Students who did not achieve a proficient score on the two previous test administrations were given another opportunity in July 2003. The highest score was reported for accountability purposes.

The performance levels corresponding to the scaled score ranges used for reporting the October 2002 and March 2003 fourth grade reading test results are shown in Table 3. A student must reach a scaled score of at least 217 on the OPT in order to be considered at or above grade level in reading.

Table 3

Performance Levels Corresponding to Scaled Score Ranges

Level	Scaled Score
Advanced	250 and above
Proficient	217 – 249
Basic	198 – 216
Below Basic	197 and below

Results

Table 4 shows the Pearson correlation coefficients among the ORF and OPT scores. The correlation analysis examined the relationship between third grade ORF for spring, fourth grade ORF for fall and spring, the fourth grade OPT for fall, and the highest score a student obtained across the three opportunities to take the OPT. Correlation coefficients demonstrate significant correlations between ORF and OPT for reading.

Table 4

Correlation Coefficients of ORF and OPT Variables

ORF/OPT Correlation	Grade 4 OPT Fall 02	Grade 4 OPT Highest Score
ORF – Grade 3	.650**	.629**
Spring 02	n = 318	n = 320
ORF – Grade 4	.646**	.654**
Fall 02	n = 355	n = 360
ORF – Grade 4	.612**	.609**
Spring 03	n = 347	n = 364

Note. ** Correlation is significant at the 0.01 level (two-tailed).

The linkage between spring of third grade DORF for students in third grade in the 2001-02 academic year and their fourth grade performance on the OPT in the fall of 2002-03 academic year is shown in Table 5. Of the students who scored at the benchmark goal of 110 or above on DORF in the spring, 172 of the 238 students (72%) scored proficient or advanced on the fall OPT. However, only 1 out of the 24 (4%) of the students who scored less than 80 on the spring

DORF was proficient on the OPT. It would appear that the third grade benchmark of 110 is sufficient for establishing a reasonable probability of proficient or advanced levels on the fall OPT.

Table 5

ORF Grade 3 Spring and Percent Passing Fall Grade 4 OPT

	DIBELS	Grade 4 OPT	
Measure	Indicator	Fall 02	
		Not Proficient	Proficient
DORF	At Risk	96%	4%
Grade 3 Spring 02	< 80 cwpm	n=23	n=1
	Some Risk	66%	34%
	80 – 109 cwpm	n=37	n=19
	Low Risk	28%	72%
	≥ 110 cwpm	n=66	n=172

The linkage between fall of fourth grade CBM ORF for students in fourth grade in the 2002-03 academic year and their performance on the fall OPT is shown in Table 6. Of the students who scored at the benchmark goal of 93 or above on CBM ORF in the fall, 186 of the 260 students (72%) scored proficient or advanced on the fall OPT. However, only 1 out of the 30 (3%) of the students who scored less than 71 on the fall CBM ORF was proficient on the OPT. It would appear that the fourth grade benchmark of 93 also is sufficient for establishing a reasonable probability of proficient or advanced levels on the fall OPT.

Table 6

ORF Grade 4 Fall and Percent Passing Fall Grade 4 OPT

	DIBELS	Grade 4 OPT	
Measure	Indicator	Fall 02	
		Not Proficient	Proficient
CBM ORF	At Risk	97%	3%
Grade 4 Fall 02	< 71 cwpm	n = 29	n = 1
	Some Risk	66%	34%
	71 – 92 cwpm	n = 43	n = 22
	Low Risk	28%	72%
	≥ 93 cwpm	n = 74	n = 186

The linkage between the fall fourth grade CBM ORF performance and the highest score on the OPT is shown in Table 7. Of the students who scored at the benchmark goal of 93 or above on CBM ORF in the fall, 233 of the 263 students (89%) scored proficient or advanced when given three attempts to pass the OPT. However, only 8 out of the 31 (26%) of the students who scored less than 71 on the fall CBM ORF were proficient on the OPT under these conditions. It would appear that the fall fourth grade benchmark of 93 is sufficient for establishing a high probability of proficient or advanced levels on the OPT when given three opportunities to pass the test. In this case, students who were at risk on ORF in the fall remained unlikely to pass the fourth grade OPT at any point during the year. The degree of their progress during the year is not reported.

Table 7

ORF Grade 4 Fall and Percent Passing with Grade 4 OPT Highest Score

Measure	DIBELS Indicator	Grade 4 OPT Highest Score	
		Not Proficient	Proficient
CBM ORF	At Risk	74%	26%
Grade 4 Fall 02	< 71 cwpm	n = 23	n = 8
	Some Risk	30%	70%
	71 – 92 cwpm	n = 20	n = 46
	Low Risk	11%	89%
	≥ 93 cwpm	n = 30	n = 233

The linkage between spring fourth grade CBM ORF performance and the highest score on the OPT is shown in Table 8. Of the students who scored at the benchmark goal of 118 or above on CBM ORF in the spring, 211 of the 235 students (90%) scored proficient or advanced when given three opportunities to attempt the OPT. Twenty-five out of the 54 (46%) of the students who scored less than 96 on the spring CBM ORF were proficient on the OPT. It would appear that the fourth grade benchmark of 118 is sufficient for establishing a high probability of proficient or advanced levels on the OPT when given intervention and multiple attempts to pass.

Table 8

ORF Grade 4 Spring and Percent Passing with Grade 4 OPT Highest Score

Measure	DIBELS Indicator	Grade 4 OPT Highest Score	
		Not Proficient	Proficient
CBM ORF	At Risk	54%	46%
Grade 4 Spring 03	< 96 cwpm	n = 29	n = 25
	Some Risk	29%	71%
	96 – 117 cwpm	n = 22	n = 53
	Low Risk	10%	90%
	≥ 118 cwpm	n = 24	n = 211

The findings of this study were confirmed with a smaller sample from one of the schools in the 2001-2002 school year.

Conclusions

The relationship between DIBELS/CBM and the Ohio Fourth Grade Reading Proficiency Test (OPT) were examined in two ways, correlations and the adequacy of DIBELS/CBM criteria as year-end goals or indicators of need for reading intervention. In general, with this sample DIBELS/CBM performance has an adequate relationship with a standardized test of reading, and benchmark goals and 'at-risk' criteria would appear valid for setting goals and deciding which students need interventions. Since DIBELS/CBM scores during fall were actually used to place students into interventions, validity analyses discussed below are confounded by having used the predictive criteria to help prevent failure on the test. Clearer conclusions may have been possible

if this biasing relationship were not present, although given earlier research, failure to intervene would have had negative ethical implications and a negative impact on student performance.

The correlations between ORF measures and the OPT reading assessment were moderately high. The current correlations were very similar to those found between ORF scores and the Oregon State Assessment and the Michigan Educational Assessment Program (Good, Simmons, and Kame'enui, 2001; McGlinchey & Hixson, 2004), higher than the correlation found with the Washington Assessment of Student Learning (Stage and Jacobsen, 2001), and lower than the correlations found in Illinois, Florida, Colorado, and North Carolina (Barger, J. 2003; Buck, J., & Torgesen, J. 2003; Good, et. al., 2001; Shaw, R., & Shaw, D. 2002; Sibley, D., Biwer, D., & Hesch, A., 2001).

One possible explanation for the lower (than a number of other states) magnitude of the correlation between ORF measures and OPT reading assessment is that the OPT reading assessment measures a variety of skills beyond the typical construct of reading. For example, correct responses often require application of critical thinking skills within a literacy context rather than just reading skills, and other items require extended written responses (both short answer and extended responses).

It should also be noted that there were some unusual patterns of student success on the OPT. In this school district, fourth grade students who scored within the proficient range or below in October were required to take the OPT again in March. Analysis of the data indicated that approximately 30% of the students who performed at the proficient level in October performed below the proficient level on the second administration in March. This sort of performance decrement may have resulted from changes in student motivation. There is heavy emphasis on students being proficient and once that was obtained in October, many students may

simply not have been motivated to do their best at the second administration. Correlations between spring DIBELS/CBM and highest OPT scores were the lowest of all comparisons, although not by a large magnitude. Even so, the patterns of correlations were similar between DIBELS/CBM at the end of grade 3, beginning of grade 4, and end of grade 4, and OPT scores in the fall and for the highest across the school year.

The DIBELS/CBM criteria were also evaluated to determine if they are valid in determining need for intervention, and as socially meaningful and valid year-end goals for individual students. Unlike the traditional evaluation of diagnostic validity, a single 'cut score' is not used with DIBELS/CBM; rather multiple cut-scores are used for different purposes. In all of the analyses conducted examining the adequacy of DIBELS/CBM as valid year-end benchmarks, the data are highly favorable and support use for individual goal setting. Students who met DIBELS/CBM benchmarks were highly likely to pass the reading OPT. False negative rates were 27% and 28% for spring/fall DIBELS/CBM and fall OPT. Such rates are not unexpected when the OPT is intended to represent end of fourth grade competence. False negative rates were around 10% when the fall or spring fourth grade ORF and highest OPT scores are analyzed. Students in the middle category ('some risk') were successful on the OPT at rates between the highest and lowest risk categories, as should be the case.

The criteria for determining at-risk status (and thus need for intervention) appeared valid when the end of third grade and beginning of fourth grade DIBELS/CBM performance were compared to fall OPT scores, and to a lesser extent when compared to the highest OPT score attained by students at any point during the fourth grade. Only 3% and 4% respectively of students scoring in the DIBELS/CBM at-risk category either in spring of the third grade or in fall of the fourth grade were successful on the fall grade four reading OPT (false positives). When

the fall DIBELS/CBM data are compared to the highest attained OPT, the false positive rate increases to 26%; however, these students were all provided reading interventions beginning in the fall and this likely explains the increase in errors. It seems better to attribute these 'errors' to the efficacy of intervention.

The much higher false positive rate for the at-risk category when spring DIBELS/CBM and highest OPT are considered are not as positive but much less important. First, spring DIBELS/CBM measures were given nearly two months after the spring administration of the OPT during which student interventions were continuing. Several months of intervention could have impacted performance on the OPT in a different manner than on the DIBELS/CBM measures. Second, the possible motivational changes discussed above and multiple test administrations may have also impacted the results. Finally, May DIBELS/CBM data are not used in the same way as earlier and validity judgments should be made reflecting different use. At the end of the year, teachers are unlikely to be making decisions to intervene with students, the purpose of the lower criteria. At times when interventions are actually planned, the at-risk criteria were highly valid.

Author Note: We would like to thank Roland Good, University of Oregon, Jennifer Desiderio, and Jim Vander Meer who reviewed this document. We would also like to thank the staff from Fairfield City Schools for their assistance with this study. Questions regarding this report can be sent to Carolyn Vander Meer at the University of Cincinnati. Her e-mail address is vandercn@email.uc.edu.

References

Barger, J. (2003). *Comparing the DIBELS oral reading fluency indicator and the North Carolina end of grade reading assessment* (Technical Report). Asheville, NC: North Carolina Teacher Academy.

Buck, J., & Torgesen, J. (2003). *The relationship between performance on a measure of oral reading fluency and performance on the Florida Comprehensive Assessment Test* (FCRR Technical Report #1) Tallahassee, FL: Florida Center for Reading Research.

Durr, W. K., & Pikulski, J. J. (1986). Houghton Mifflin Reading. Boston: Houghton Mifflin Company.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Waltz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology review, 22*, 27-48.

Good, R. H., Gruba, J., & Kaminski, R. A. (2001). Best Practices in Using Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in an Outcomes-Driven Model. In A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology IV,* (pp. 679-700). Washington, DC: National Association of School Psychologists.

Good, R. H., & Jefferson, G. (1998). Contemporary perspectives on curriculum-based measurement validity. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 61-88) New York: Guilford Press.

Good, R. H., & Kaminski, R. A. (Eds.). (2002*). Dynamic Indicators of Basic Early Literacy Skills (6th ed.)*. Eugene, OR: Institute for the Development of Education Achievement. Available: http://dibels.uoregon.edu/

Good, R. H., Kaminski, R. A., & Dill, S. (2002). DIBELS Oral Reading Fluency. In R. H.

Good & R. H. Kaminski (Eds.), *Dynamic Indicators of Basic Early Literacy Skills (6th ed.)*. Eugene, OR: Institute for the Development of Educational Achievement.

Good, R. H., Kaminski, R. A., Shinn, M., Bratten, J., Shinn, M., Laimon, L. (2003*). Technical adequacy and decision making utility of DIBELS* (Technical Report). Manuscript in preparation. Eugene, OR: University of Oregon.

Good, R. H., III, Simmons, D. C. & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*, 257-288.

Hasbrouck, J. E., and Tindal, G. (1992, Spring). Curriculum-based oral reading fluency norms for students in grades 2 through 5. *Teaching Exceptional Children, 24*, 41- 44.

Kaminski, R. A., & Good, R. H., III. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25*, 215-227.

Marston, D. (1989). Curriculum-based measurement: What is it and why do it? In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18- 78), New York: Guilford Press.

McGlinchey, M. T. & Hixson, M. D., (2004). Using Curriculum-Based Measurement to Predict Performance on State Assessments in Reading. *School Psychology Review, 33*, 193-203.

Ohio Department of Education, Statistical Summary of March 2003 Grade 4 and 6 Proficiency Tests. www.ode.state.oh.us/proficiency/ retrieved from proficiency and testing home page 7/15/04.

Shaw, R., & Shaw, D. (2002). *DIBELS oral reading fluency-based indicators of third grade reading skills for Colorado State Assessment Program (CSAP).* (Technical Report).

Eugene, OR: University of Oregon.

Shinn, M. R. (Ed.). (1989*). Curriculum-Based Measurement: Assessing Special Children*. New York: Guilford Press.

Shinn, M. R. (1997). *Instructional Decision Making Using Curriculum-Based Measurement*. Unpublished workshop materials.

Shinn, M. R. (Ed.) (1998). *Advanced Applications of Curriculum-Based Measurement*. New York: The Guilford Press.

Shinn, M. R., Good, R. H., Knutson, N., Tilly, W. D., & Collins, V. L. (1992). Curriculum-based measurement reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review, 21*, 459-479.

Sibley, D., Biwer, D., & Hesch, A. (2001). Establishing curriculum-based measurement oral reading fluency performance standards to predict success on local and state tests of reading achievement. Unpublished data. Arlington Heights, IL: Arlington Heights School District 25.

Stage, S. A., Jacobsen, M. D., (2001). Predicting student success on state-mandated performance-based assessment using oral reading fluency. *School Psychology Review, 30*(3), 407-419.

Tindall, G., Marston, D., & Deno, S. L. (1983). *The reliability of direct and repeated measurement* (Research Rep. No. 109). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.

University of Oregon (2004). DIBELS benchmark levels. Retrieved from http://dibels.uoregon.edu/benchmark.php.