Development of the Standard Reading Passages

Introduction

There are many reading tests currently available to educators, although, very few of these tests measure growth - particularly growth in reading fluency[1]. Experts in the field of reading, however, long have emphasized the significance of fluency. The Commission on Reading (1985), for example, stated in its A Nation of Readers that, "readers must be able to decode words quickly and accurately so that this process can coordinate fluidly with the process of constructing the meaning of the text" (p. 11). Further, the Commission on Reading stated that "more comprehensive assessments of reading are needed. Standardized tests do not provide a deep assessment of reading comprehension and should be supplemented with observations of reading fluency..." (p. 101).

More recently, the role of fluency has received considerable attention. In its review of research on reading instruction, the National Reading Panel (2000) emphasized reading fluency has a significant place in literacy programs:

> "Fluent readers are able to read orally with speed, accuracy, and proper expression. Fluency is one of several critical factors necessary for reading comprehension" (p. 11).

Further, a publication written by the Center for the Improvement of Early Reading Achievement (CIERA) and funded by the National Institute for Literacy (NIFL) states,

> "Monitoring your students' progress in reading fluency will help you determine the effectiveness of your instruction and set instructional goals. Also, seeing their fluency growth reflected in the graphs you keep can motivate students" (2001. p. 30).

Typically, standardized tests of reading achievement provide the educator with static assessments of components of reading such as "comprehension", "word attack", "vocabulary", "word study skills", "consonant and vowel sounds", and so on. In general, these tests have provided educators with scores for each of those components that describe student performance relative to the norm group. While those scores give a picture of a student relative to peers at a moment in time, norm-referenced scores cannot provide direct information on growth in performance across time. For example, a student who is at the 20th percentile in reading in the beginning of one school year who grows as much as his peers will still be at the 20th percentile at the beginning of the next year. Measuring growth simply has not been a goal or purpose of norm-referenced test development. This failure to clearly and directly describe a student's growth across time is a particular problem for teachers who are trying to evaluate their success in teaching a child to read.

A second important problem for using most tests is that they are insensitive to growth over relatively short time periods. It is not unusual to find that a typical norm-referenced achievement test will not reliably reveal change in a student's performance for six months or a year. For teachers who are trying to be responsive to their students' progress and make changes in instruction during the course of the school year, this is a serious problem. Teachers need assessment procedures that they can use routinely to monitor the growth of their students and make frequent instructional changes when progress rates are unsatisfactory.

Finally, standardized norm-referenced achievement tests historically have not measured the fluency of a child's performance. In general, these tests have focused on accuracy of responding without regard for the time it

[1] The term "fluency" has been defined in many different ways. We are using the term here in its most common sense; that is, the accuracy and speed of reading text.

takes a student to respond. In part, this has occurred because test developers have been insufficiently attentive to the key role that speed (or latency) of responding plays in cognitive processing and comprehension in reading. Another reason for this limitation of traditional achievement test is that they have been designed for group administration, and it is easier to design tests for group administration that do not require that performance be timed. One result of this limitation is that teachers have been encouraged to focus solely on accuracy of responding without considering the rate of student responding. That is unfortunate because, as we have seen in connection with previous recommendations from panels of reading experts, considerable evidence now exists that a child's success in learning to read will be reflected in the fluency with which he reads text.

The *Standard Reading Passages* have been created to capitalize on the emergent knowledge regarding the importance of oral reading fluency as an indicator of growth in general reading proficiency (Fuchs, et al, 2001). The *Standard Reading Passages* are the first passages written for educators to use as a basis for evaluating growth in student's oral reading fluency. Originally distributed as *The Test of Oral Reading Fluency (TORF),* the *Standard Reading Passages* are a revised and updated set of passages that can be used to assess individual student growth and reveal differences between students in reading proficiency.

Purposes of the Standard Reading Passages

The *Standard Reading Passages* have been designed for two specific educational purposes. First, the *Standard Reading Passages* can be used as a screening tool to determine whether students are at risk for developing a reading problem. Those students can be flagged for closer scrutiny, for alternative instruction, or for more intensive assessment. A second use of the *Standard Reading Passages* is to provide a means for regular and frequent progress monitoring to assure the achievement of important benchmarks. For these reasons, the *Standard Reading Passages* consist of two parts.

Screening

Screening is a process by which a group of students is quickly appraised to get a picture of who is doing well and who is doing poorly. The word screening comes from the use of a screen to filter out particles of different sizes. The Screening section of the *Standard Reading Passages* does just that for the area of reading.

With the *Standard Reading Passages* educators are provided with a set of passages calibrated for each grade level from first through sixth grade. A student will read the passages and the educator scores each sample for accuracy and speed. To determine if the student is discrepant from peers, the *Standard Reading Passages* can also be administered to a random set of peers at either the same grade level. The resulting distribution of reading scores can then be used to identify those students whose discrepancy in performance is indicative of potential reading problems.

Progress Monitoring

Current concerns over the large number of students failing to achieve satisfactory levels of literacy have increased the emphasis on the role of monitoring student growth to assure critical educational outcomes. This means that educators are increasingly being called upon to evaluate the effectiveness of a student's instructional program and to revise that program if it is not producing satisfactory progress toward important benchmarks. The importance of teacher's responsiveness to student progress during the course of the school year is not simply established through legal mandates. Considerable evidence exists that teachers' use of progress monitoring or "formative assessment" is an essential ingredient for effective instruction. Indeed, some have argued that there is

> "…no other way [than formative assessment] of raising standards for which such a strong prima facie case can be made." (Black & William, 1998)

The Progress Monitoring section of the *Standard Reading Passages* was designed for use in appraising the rate of improvement of all students learning to read English.

Theoretical Premises Underlying the Standard Reading Passages

Direct Measurement

A primary assumption on which the *Standard Reading Passages* are based is that assessment should be accomplished through directly observing samples of student's academic skill performance. The observation of performance samples differs from the approach used in most standardized tests where students typically respond to multiple choice items or "yes-no" questions from which inferences about a student's knowledge are made. Certainly, such a test format has provided educators with useful information about pupil performance; however, the precision of this model for measuring what the child actually knows or can do has been questioned (Elliott & Piersel, 1982; Ysseldyke & Marston, 1982). As one brief example, Tuinman (1971) demonstrated with several measures of reading comprehension that multiple choice items can be answered correctly without reading the corresponding passage. Lovitt (1967) argued effectively that the validity of assessment was significantly improved with an emphasis upon the direct measurement of academic performance of interest. Thus, the educator who wishes to determine the extent of a child's reading skill should directly observe the student reading text. To facilitate this observation, one obvious approach is to have the child read aloud.

Repeated Measurement

A second important feature of the *Standard Reading Passages* is the repeated measurement of pupil performance. Repeated measurement is the primary strategy for progress monitoring and is described in that section of this manual. From the measurement theorist's perspective, repeated behavioral sampling enhances any assessment system because it improves reliability (Nunnally, 1978; Thorndike & Hagen, 1978). Most standardized achievement tests can be administered once, perhaps twice, in an academic year. If we are interested in measuring student growth frequently during the school year, we will be at a loss with the standardized test approach. As Deno (1985) stated "A teacher who is attempting to make daily and weekly instructional effectiveness decisions needs frequent information that reflects whether or not a student's performance is increasing within the span of those days and weeks. Typical achievement tests were not constructed to provide information on short-term growth and indeed they do not" (p. 225).

Analysis of Performance Graphs

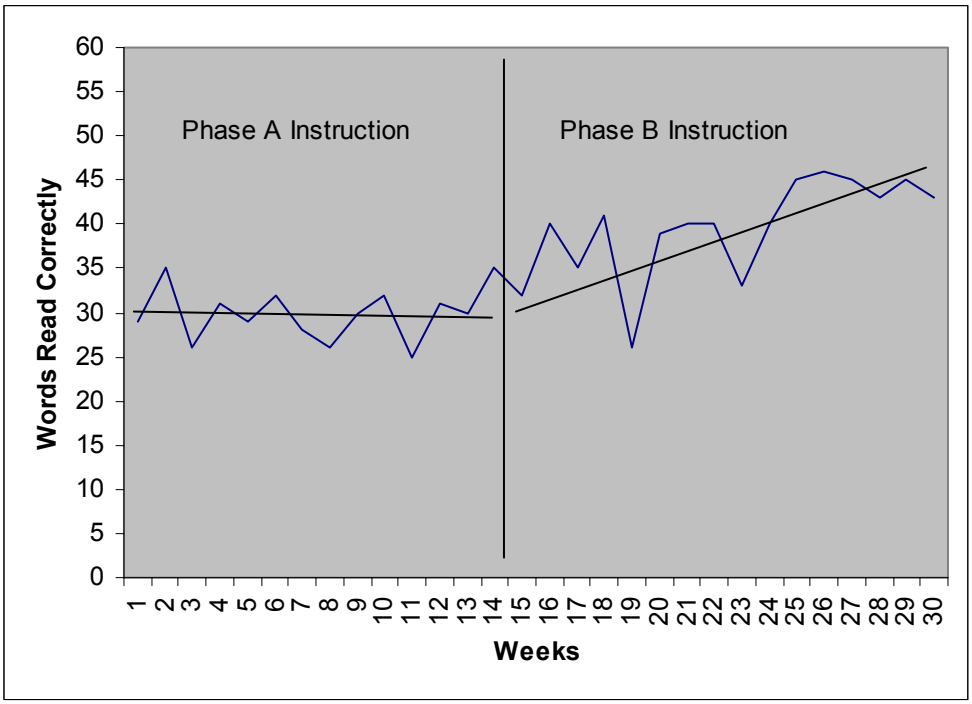
A third major element of progress monitoring with the *Standard Reading Passages* is analysis of the graphed reading data. While direct and repeated measurement of pupil performance are necessary conditions for successful implementation of the *Standard Reading Passages,* they are not sufficient to assure the improved achievement outcomes. Essential to the process is the analysis of graphed student learning rates in relation to modifications in educational programs.

Research by Brandstetter and Merz (1978) demonstrated that merely collecting reading samples was not enough. The significant reading gains that can be obtained when student progress is repeatedly measured are produced when the resulting data are charted and analyzed. Fuchs, Deno, and Mirkin (1984) compared the progress of students whose reading performance was repeatedly measured, graphed and programs modified in response to the data. They found that repeated measurement and graphing of reading fluency data was related to greater gains in reading. Further, Fuchs and Fuchs (1986) meta-analysis of research on formative evaluation revealed that student achievement was approximately .7 standard deviation units greater when teachers learned to use graphed data to evaluate and modify their instruction.

An illustration of how reading performance graphs are analyzed is shown in Figure 1 below. In this example the number of words the student reads correctly during a one minute timing on equivalent passages is graphed weekly. As can be seen, the learning rate of the student is markedly different in Phase A and Phase B instruction. The instructional strategy used in Phase A creates very little growth and has a slope of about .1 correct word increase per week. In contrast, Phase B instruction appears to significantly benefit the student. The value of the slope line in Phase B is approximately 1.2 correct word increase per week.

Graphing and evaluating student reading performance in this fashion provides the teacher with two benefits. First, the graphed data demonstrate the extent to which the student is improving. Second, the evaluation system provides the teacher with feedback on the effectiveness of his/her instruction.

Figure 1. An example of graph analysis with direct and measurement of reading fluency



Technical Adequacy of the Standard Reading Passages

All assessment procedures used in education must conform to certain standards that assure the consumer that the product they are using is satisfactory. Guidelines set out by the American Psychological Association, the American Educational Research Association and the National Council for Measurement in Education (1985) provide the public with such standards. According to these APA Test Standards, a test must have evidence of validity and reliability, and adequate norms. This section reviews research data on the validity and reliability of the *Standard Reading Passages.* Norms are not provided for the *Standard Reading Passages*; however, benchmark guides are provided as an alternative to increase the usefulness of the tools for decision-making.

Validity

The procedures used for measuring student reading growth with *the Standard Reading Passages* are those developed originally for Curriculum-Based Measurement (CBM) (Deno, 1985), and more recently described as General Outcome Measurement (GOM) (Fuchs & Deno, 1994). The technical adequacy of the procedures used in CBM has been extensively examined for more than twenty-five years (Deno, 2003) and they continue to be the source of important empirical studies in assessment (School Psychology Review, 2003). That research history on establishing the validity of oral reading fluency as an indicator of reading growth began with an important study by Deno, Mirkin and Chiang (1982). These researchers identified five measures of reading that could potentially be used to index pupil reading proficiency: (1) asking the student to read aloud stories from the pupil's basal reader (Passage Reading); (2) asking the student to read aloud lists of words randomly selected from the pupil's basal reader (Isolated Word Lists); (3) asking the student to read aloud words underlined in a story from his/her basal reader (Reading in Context); (4) asking the student to supply words that had been deleted from stories from his/her basal reader (Cloze Comprehension Procedure); and (5) asking the student to give the meaning of words selected from the basal reader story (Word Meaning). These reading measures were then correlated with widely used standardized tests of reading. The criterion measures used in this initial criterion-related validity study were the *Stanford Diagnostic Reading Test* (Karlsen). Madden, & Gardner, 1975), the *Woodcock Reading Mastery Test* (Woodcock, 1973) and the Reading Comprehension subtest from the *Peabody Individual Achievement Test* (Dunn & Markwardt, 1970). Deno et al. (1982) found that listening to a child read aloud from his/her basal reader for one minute correlated strongly with the standardized achievement tests. Validity coefficients in the study ranged from .73 to .93.

Further confirmation of this finding was provided by Fuchs and Deno (1981), Fuchs, Tindal, Shinn, Fuchs, Deno and Germann (1983), Fuchs, Tindal and Deno (1984), Marston and Magnusson (1985) and Fuchs, Fuchs, and Maxwell, (1988) who demonstrated that oral reading fluency correlated significantly with other standardized tests of reading. In the Fuchs and Deno (1981) study, criterion tests included the Woodcock Reading Mastery Test (Woodcock, 1973) and teacher judgment. The Ginn 720 Mastery Total Test was used as a criterion in the Fuchs, et al. (1983) study. Fuchs, Tindal and Deno (1984) utilized a Cloze comprehension task and a word meaning test as a criterion for validity. Marston and Magnusson (1985) contrasted words read correctly from basal passages by 3rd graders with their performance on *Science Research Associates Achievement Series* (Naslund, Thorpe, & Lefever, 1978), the *Stanford Achievement Test* (Madden, Gardner, Rudman, Karlsen, & Merwin, 1973) and student placement in the basal reading series, in this case *Ginn 720* (Clymer & Fenn, 1979). Finally, Fuchs, Fuchs, and Maxwell (1988) examined oral reading fluency in relation to the Comprehension subtest of the Stanford Achievement Test and other commonly used informal measures of comprehension (Question Answering, Story Retelling, and Cloze). The range of validity coefficients found in these studies is summarized in Table 1 (Tindal and Marston (1986). One of the most remarkable findings contained in Table 1 is the set of correlations obtained by Fuchs, Fuchs, and Maxwell. In that study, the mean correlation of oral reading fluency with the various comprehension measures was higher than the correlations of those measures to one another.

Shinn and his associates (Shinn, et al, 1992) conducted a confirmatory factor analysis of the role played by oral reading fluency in general reading achievement. Their results revealed that oral reading fluency accounted for virtually all of the predictable variation in achievement. They concluded that the results were sufficiently persuasive that the oral reading measure of CBM could be considered to possess construct validity as a measure of general reading proficiency.

Table 1. A summary of validity studies examining the relationship between words read correctly and criterion measures.

Study	*Grades*	*Coefficients*
Deno, Mirkin, & Chiang (1982)	1-5	.73-.93
Fuchs & Deno (1981)	1-6	.82-.86
Fuchs, Tindal, Shinn, Fuchs, Deno, & Germann (1983)	5	.52-.85
Fuchs, Tindal, & Deno (1984)	3-6	.60-.84
Marston & Magnusson (1985)	3	.59-.90
Fuchs, Fuchs, & Maxwell (1988)	4-8	.80-.91

In addition, Deno, Marston, Shinn and Tindal (1982) established the discriminative validity of oral reading when they reliably differentiated LD and non-LD samples of students. This finding was replicated by Shinn and Marston (1985) who found that words read aloud differentiated regular education student, pupils served in Chapter I and mildly handicapped students with learning difficulties. Marston, Tindal, and Deno (1984) demonstrated that words correctly read aloud predicted LD classification as well as traditional measures of aptitude /achievement discrepancy. Moreover, the direct measures of student reading required significantly smaller expenditures in terms of cost and time. In addition, Marston, Mirkin and Deno (1984) demonstrated that for screening purposes the reading fluency measure was as accurate as traditional measures and added objectivity to teacher referrals.

The validation of the words read correct was also demonstrated in a series of studies on the topic of sensitivity to growth. Reasoning that a valid measure of reading must show growth as student skills improve, several investigations were initiated. A cross-sectional study of oral reading fluency across grades one through six in a sample of 550 elementary students showed reliable gains (Deno, Marston, Mirkin, Lowry, Sindelar & Jenkins, 1980). Marston, Fuchs, and Deno (1986) examined the short term reading progress of students across 10-week and 16-week intervals with both standardized reading tests and words read correctly from grade level passages. While both approaches identified student improvement, the reading fluency measure (1) delineated greater growth in the reading performance of students, and (2) correlated much higher with teacher perceptions of individual student improvement.

A validity study of the Standard Reading Passages was conducted by Ikeda, Gruba, Dunga (2000). These researchers used the Standard Reading Passages (at that time known as the Test of Reading Fluency) to measure the reading performance of 112 students in the Fall and Spring of 2nd and 3rd grade, and in the Fall of 4th grade. These students were then administered the Iowa Test of Basic Skills (ITBS) 4th Grade Reading Comprehension test. In addition, discriminant analyses were run for each CBM administration to examine how well the Standard Reading Passages predicted performance on the ITBS at 4th grade. Intercorrelations between the administrations of the Standard Reading Passages were high, with a range of .72 to .92 and a median of .87. As shown in the table below, all correlations between performance on the Standard Reading Passages and the ITBS were statistically significant at the .01 level. In addition, the discriminant function analyses showed performance on the Standard Reading Passages successfully predicted proficiency on the ITBS with accuracy rates ranging from 70% to 77%.

Table 2. Correlations, P Values and Classification Rates for the Standard Reading Passages predicting ITBS reading comprehension*

Time of Test Administration	*Correlation*	*P Value*	*Classification Rate*
Fall 2nd Grade	.66	.01	72%
Spring 2nd Grade	.65	.01	72%
Fall 3rd Grade	.65	.01	71%
Spring 3rd Grade	.61	.01	74%
Fall 4th Grade	.71	.01	80%

* Data abstracted from Ikeda, Gruba, & Dunga, (2000).

Good, Simmons, and Kame'enui (2001) studied the validity of the Standard Reading Passages, also known at the Test of Reading Fluency in that investigation. They report that Good & Jefferson (1998) examined criterion-validity in eight studies and found coefficients ranging from .52 - .91. Further, Good, Simmons, and Kame'enui (2001) studied the predictive validity of *Standard Reading Passages* in the Spring of 1st, 2nd, and 3rd grades **and** 3rd grade performance on the Oregon Statewide Assessment - Reading/Literature. These researchers found a correlation of .82 between Spring Grade 1 and Grade 2 administrations of the *Standard Reading Passages*. The correlation between the *Standard Reading Passages* in Spring Grade 3 and the Oregon State Assessment at Grade 3 was .67.

Further investigations of the relationship between the Standard Reading Passages and high stakes testing were conducted by Deno (2003). In a series of studies the predictive validity of oral reading – words correct and maze reading – correct word choices was examined. In these studies the criterion variable was the MAT-7 reading test and the Minnesota Comprehensive Assessment (MCA) of reading. These analyses were conducted at both 3rd and 5th. A review of the inter-correlations between oral reading and the corrected MAZE scores for the Fall, Winter, and Spring administrations in the 5th grade ranged from .67 to .82 ($p<.01$). For a group of 5th graders, the correlation between Words Read Correctly and MAT-7 Reading was .77 and between Words Read Correctly and the MCA Reading test the coefficient was also .77. In a study of 3rd and 5th graders administered maze probes from the Standard Reading Passages, the correlation for 3rd grade students was .66 between maze and MAT-7 and .66 for 5th graders. For 5th graders the correlations were .65 with the MAT-7 and .65 with the MCA.

Reliability

A second major component of the APA Test Standards is reliability. Reliability is the extent to which we get stable and consistent results from our testing. As outlined in the guidelines, there are several important types of reliability, including: test-retest reliability, alternate form reliability, and inter-rater agreement.

Tindal, Marston, and Deno (1983) studied the test-retest reliability of elementary aged students and determined that passage reading scores were stable across time. The test-retest coefficients ranged from .92 to .97. In the same study the alternate-form reliability of different reading passages drawn from the same level was examined. Alternate-form reliabilities ranged from .89 to .94. Finally, Tindal, Marston, and Deno (1983) examined the extent to which the scores of trained observers on reading passages agreed with one another. Their data indicated that observer reliability was very high with a median coefficient of agreement of .99.

Norms

The intention of the authors, in designing the screening portion of the *Standard Reading Passages,* was that users would use the test at the classroom or school level and develop their own local normative database. In this way users will become more familiar with typical performance within their own classrooms and schools and use the test for screening in these contexts rather than as an identification measure in which nationally-based norm-referenced scores are emphasized.

References

American Psychological Association, American Educational Research Association, and National Council on Measurement in Education. *(1985). Standards for educational and psychological tests.* Washington, D.C.: American Psychological Association.

Black, P. & Wiliam, D. (1998) Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80, (2)* Pp 139-148.

Brandstetter, G., & Merz, C. (1978). Charting scores in precision teaching for skill acquisition. *Exceptional Children, 45, 42-48.*

Center for the Improvement of Early Reading Achievement (2001). Put reading first: The research building blocks for teaching children to read. Washington, DC: The Partnership for Reading.

Clymer, T., & Fenn, T. *(1979). Reading 720: Rainbow Edition.* Lexington, MA: Ginn.

Commission on Reading. *(1985). A nation of readers: The report of the commission on reading.* Washington, D. C.: The National Institute of Education.

Deno, S. L. (2003). A school-wide model for progress monitoring and predicting success on high stakes outcomes. Presentation at annual Conference of the Council for Exceptional Children, Seattle, WA.

Deno, S.L.(2003) Developments in Curriculum-based Measurement. Accepted for publication in the Journal of Special Education

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative*. Exceptional Children, 52*, 219-232.

Deno, S. L., Mirkin, P. K., & Chiang, B. *(1982).* Identifying valid measures of reading. *Exceptional Children, 49, 36-45.*

Deno, S. L., Marston, D., Mirkin, P., Lowry, L., Sindelar, P. & Jenkins, J. *(1982). The use of standard tasks to measure achievement in reading, spelling, and written expression: A normative and developmental study* (Research Report No. *87).* Minneapolis, MN: Institute for Research on Learning Disabilities, University of Minnesota.

Deno, S. L., Marston, D., Shinn, M., & Tindal, J. *(1983).* Oral reading fluency: A simple datum for scaling reading disability. Topics *in Learning and Learning Disabilities, 2(4), 53-59.*

Dunn, E. K. & Eckert, T. L. (2002) Curriculum-based measurement in reading: A comparison similar versus challenging material. *School Psychology Quarterly 17*, (1) pp. 24-

Dunn, L. M. & Markwardt, F. C. *(1970). Peabody individual achievement test.* Circle Pines, MN: American Guidance Service.

Elliott, S. N. & Piersel, W. C. *(1982).* Direct assessment of reading skills: An approach which links assessment to intervention. *School Psychology Review, 11, 267-280.*

Fuchs, L.S., & Deno, S.L. (l994). Must instructionally useful performance assessment be based in the curriculum? Exceptional Children, 61, 15-24

Fuchs, L. & Deno, S. L. *(1981). The relationship between curriculum- based mastery measures and standardized achievement tests in reading* (Research Report No. *57).* Minneapolis, MN: University of

Minnesota, Institute for Research on Learning Disabilities. (ERIC Document Reproduction Service No. ED *212 662)*

Fuchs, L., Deno, S. L., & Mirkin, P. K. (1984). The effects of frequent curriculum based measurement and evaluation on pedagogy, student achievement, and student awareness of learning. *American Educational Research Journal,* 21, 449-460.

Fuchs, L. S. & Fuchs, D. (1986). Effects of systematic formative evaluation: A metaanalysis. *Exceptional Children, 53,* 199-208.

Fuchs, L.S., Fuchs, D., Hosp, M., & 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.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. Scientific Studies of Reading, 5(3), 239-256.

Fuchs, L. S., Fuchs, D., & Maxwell, S. (1988). The validity of informal reading measures. *Remedial and Special Education, 9* (2), 20-28.

Fuchs, L. S., Tindal, G., & Deno, S. L. (1984). Methodological issues in curriculum based reading assessment. *Diagnostique, 9,* 191-207.

Fuchs, L. S., Tindal, G., Shinn, M. R., Fuchs, Deno, S., & Germann, G. (1983). *Technical adequacy of basal readers' mastery tests: Ginn 720 series* (Research Report No. 122). Minneapolis, MN: University of Minnesota Institute for Research on Learning Disabilities.

Good, R. H., 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 (3), 257-288.

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

Ikeda, M. J., Gruba, G. G., & Dungan, S. G. (2000.). Using Curriculum-Based Measures to establish "Benchmarks" predictive of reading comprehension. Johnston, IA: Heartland ADEA 11.

Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C.A. & Deno, S. L. (accepted for publication) Sources of Individual Differences in Reading Comprehension and Reading Fluency. Journal of Educational Psychology.

Karlsen, B., Madden, R., & Gardner, E. F. (1975). *Stanford diagnostic reading test.* New York: Harcourt Brace Jovanovich.

Lovitt, T. C. (1967). Assessment of children with learning disabilities. *Exceptional Children, 34,* 233-239.

Madden, R., Gardner, E., Rudman, H., Karlsen, B., & Merwin, J. (1973). *Stanford achievement test.* New York: Harcourt Brace Jovanovich.

Marston, D., Fuchs, L., & Deno, S. L. (1986). Measuring pupil progress: A comparison of standardized achievement tests and curriculum-related measures. *Diagnostique, 11,* 77-90.

Marston, D. & Magnusson, D. (1985). Implementing curriculum-based measurement in special and regular education settings. *Exceptional Children, 52,* 266-276.

Marston, D., Mirkin, P. K. & Deno, S. L. (1984). Curriculum-based measurement: An alternative to traditional screening, referral, and identification. *Journal of Special Education, 18,* 109-118.

Marston, D., Tindal, G., & Deno, S. L. (1984). Eligibility for learning disability services: A direct and repeated measurement approach. *Exceptional Children, 50* (6), 554-555.

National Reading Panel. (2000). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. City: NICHD; Retrieved from the World Wide Web at http://www.nichd.nih.gov/publications

Naslund, R. A., Thorpe, L. P., Lefever, D. W. (1978). *SRA Achievement Series.* Chicago: Science Research Associates.

Nunnally, J. (1978). *Psychometric theory.* New York: McGraw-Hill.

Salvia, J. & Ysseldyke, J. E. (1987). *Assessment in special and remedial education* (4th Ed.). Boston: Houghton-Mifflin.

Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children.* New York: 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.

Shinn, M. R. & Marston, D. (1985). Assessing mildly handicapped, low-achieving, and regular education students: A curriculum-based approach. *Remedial and Special Education, 6* (2) 31-45.

Thorndike, R., & Hagen, E. (1978). *Measurement and evaluation in psychology and education* (2nd Ed.). New York: Wiley.

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

Tuinman, J. J. (197 1). Asking reading-dependent questions. *Journal of Reading, 14,* 289292.

Woodcock, R. W. (1973). *Woodcock reading mastery test.* Circle Pines, MN: American Guidance Service.

Ysseldyke, J. E. & Marston, D. (1982). A critical analysis of standardized reading tests. *School Psychology Review, 11,* 257-266.