4-1-1990

Teacher Ratings as Predictors of Reading Achievement on Curriculum-Based and a Standardized Oral Reading Test

Harva H. Paul
University of Nebraska at Omaha

TEACHER RATINGS AS PREDICTORS OF READING ACHIEVEMENT ON CURRICULUM-BASED AND A STANDARDIZED ORAL READING TEST

A Field Project

Presented to the

Department of Psychology

and the

Faculty of the Graduate College

University of Nebraska

In Partial Fulfillment

of the Requirements for the Degree

Education Specialist in School Psychology

University of Nebraska at Omaha

by

Harva H. Paul

April 1990

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

FIELD PROJECT ACCEPTANCE

Acceptance for the faculty of the Graduate College, University of Nebraska, in partial fulfillment of the requirements for the degree Education Specialist in School Psychology, University of Nebraska at Omaha.

Committee

Name	Department
[signature]	Special Education
[signature]	Psychology

[signature]
Chairman

4/19/90
Date

Abstract

Teacher ratings as predictors of student reading speed and accuracy using curriculum-based assessment (CBA) probes and a standardized oral reading measure were evaluated. One hundred and nine third graders read orally three CBA probes from their classroom reader and the third grade level samples from the Gray Oral Reading Test. Teachers independently rated students on speed of reading, errors made, and overall reading skills. Ratings were compared with actual performance. Teacher ratings of overall reading skills best predicted performance on the CBA probes, while ratings of speed of reading and errors best predicted student outcomes on the Gray passages. Teachers were surprisingly inaccurate in predicting number of errors made on the CBA probes, but teachers were accurate in predicting students' reading speed.

TABLE OF CONTENTS

Teacher Ratings a Predictors of Reading Achievement on Curriculum-Based and a Standardized Oral Reading Test

Since assessment of student skills comprises a significant portion of our daily activities, it is the school psychologist's responsibility to provide accurate and relevant information about a child experiencing learning problems. The content validity of measures used in an evaluation is a critical factor to be considered before educational decisions can be made. As Leinhardt & Seewald (1981) argued, there must be significant overlap between what is taught and what is tested in order to measure the learning process more directly.

Content validity is a principal concern since research indicates achievement tests used to measure reading skills clearly show biases in favor of some curriculums over others (Jenkins & Pany, 1978; Webster, McInnis & Craver, 1986). A student receiving instruction in a curriculum that has little overlap with the standardized achievement test administered is likely to score lower than a student receiving instruction in a curriculum with substantial overlap. Non-curriculum based achievement tests have not only been shown to underestimate school achievement, but in some cases, overestimate academic skills (Idol-Maestas, Lloyd & Lilly, 1981; Neville & Hamm, 1985). Therefore, a factor in students' test performance is the degree of overlap.

The issue of content validity appears to be solved through the development of curriculum-based assessment procedures.

Curriculum-based assessment or CBA utilizes the students' own curriculum in assessing skills. CBA has been gaining popularity and has been described as the "emerging alternative" for evaluation of academic achievement (Deno, 1985). Curriculum-based assessment has distinct advantages over traditional standardized measures:

1. CBA resolves the question of content validity. CBA procedures have been developed for reading, spelling, math, and written language. Testing materials are taken directly from classroom texts (e.g., reading probes) or based on curriculum objectives (e.g., math probes).

2. CBA allows for the adjustment of individual educational programs based on outcome data as mandated by Public Law 94-142. CBA provides the means for frequent, continuing monitoring of individual student skills (Germann & Tindal, 1985). CBA also provides new indexes of progress including slopes of improvement and learning rate (Marston & Magnusson, 1985).

3. The use of CBA has been found to produce positive educational effects. Fuchs, Deno & Mirkin (1984) discovered teachers using CBA's frequent monitoring system were more realistic regarding student expectations. Classrooms implementing CBA were found to produce increased student performance and greater student awareness of individual goals and progress.

4. CBA compares favorably with teacher referral for academic problems and is potentially useful as an objective screening or identification measure. A study by Marston, Mirkin, & Deno, (1984) indicated CBA did not result in fewer referrals, however, teacher

referrals were made without the influence of common biases, i.e., race, sex, socio-economic class.

5. CBA lends itself beautifully to peer-referencing through the establishment of local norms. Local norms have the advantage of homogeneity over national norms (Angoff, 1971; Elliott & Bretzing, 1980; Kamphaus & Lozano, 1984). CBA local norms reflect a true comparison of a student with peers since classmates experience the same learning opportunities.

6. CBA is cost effective, time efficient, and easy to administer (Shinn, 1988). CBA requires only a few minutes and uses materials readily available in the school. Administration and scoring procedures are easy to learn. Paraprofessionals or students themselves can be trained in CBA monitoring, thus providing teacher flexibility.

A question often asked about CBA is how can such a simple and brief measure of reading skills be valid? Deno, Mirkin and Chiang (1982) found that reading aloud from a basal reader is correlated with performance on standardized reading tests. Their data also indicated that correct performance is a more valid measure of reading than error performance. This finding contradicts the premise of standardized oral reading tests which emphasize the number and types of reading errors made. In CBA errors become an important variable only when a student can not read any words correctly (Deno, et.al. 1982). Lovitt and Fantasia (1980) compared CBA with the Diagnostic Reading Scales and found both measures correlated similarly with the number of basal readers

passed per year by students. The use of oral reading fluency as a valid measure of general reading achievement has also been supported by research finding high correlations between fluency and published reading measures (Fuchs & Deno, 1981; Fuchs, Fuchs & Deno, 1982).

CBA has been found to correlate well with teacher ratings of students' academic skills. Results in a 1985 study by Marston & Magnusson comparing teacher ratings of student reading skills and rates of reading revealed high correlations. The Iowa Learning Disabilities Research Consortium (1989) reported teacher ratings to be valid in determining a student's need for special education. CBA, using words read correctly per minute, was far superior in determining students needing special education than word error measures. Studies such as this suggest teacher ratings and CBA are valid methods of identifying students with true academic problems as opposed to low achievers.

In evaluating the content validity of CBA and standardized tests the amount of overlap with the child's classroom curriculum is an important issue. Since CBA utilizes the student's actual reading text, it would stand to reason that the content validity of CBA is higher than that of a standardized test. Previously cited research revealed high correlations between teacher ratings of academic skills and CBA. Teacher ratings also correlate well with standardized test scores (Shapley & Edgar, 1986; Colardarci, 1986). It is reasonable to assume that teachers should be able to predict performance in the classroom curriculum better than performance

when a standardized test is used. Thus the hypothesis generated is that teacher ratings of student reading skills will correlate higher with students' performance on CBA measures than performance on a standardized measure of oral reading. Speed of reading and errors made served as the dependent variables.

Method

Subjects

Letters explaining the project and requesting parental consent for participation were sent to all parents of third grade students in the Glenwood Community schools, Glenwood, Iowa. Parents of 118 of the 132 students responded. Of these 118 students nine students' parents requested their child not participate, leaving 109 students who were involved in the study. Of the 109 students there were 53 males and 56 females. Both of the elementary schools in the Glenwood system were about equally represented in the study. Eight participants were receiving reading instruction in special education programs and seven participants received reading instruction through a talented and gifted program. All remaining participants received reading instruction in the regular third grade curriculum.

Materials

Three reading probes were drawn from the third grade reader, Catching Glimpses which is published by the American Book Company. This reader was chosen following interviews with teachers to determine where most third graders are expected to be at the end of the year. The majority of the third grade was currently in

Catching Glimpses and expected to finish it by the end of the year. Students not in this reader were generally in an accelerated reading program or a special education program.

Reading probes were constructed using the procedures outlined by Shapiro and Lentz (1987). A probe consisted of a passage of 200 to 250 words taken directly from the reader. Passages containing dialogue, more than two foreign or unusual words were eliminated. Poetry and plays were not used. Type style was preserved; however, pictures were eliminated as it was felt they could provide cues to the context and therefore, some of the vocabulary. Probes were drawn from the first third, middle third, and final third of the reader. Two copies were made of each passage, one for the child to read and one for the examiner to record errors and number of words read. Presentation of the passages was random.

The third grade level reading passages from Forms A, B, and C of the Gray Oral Reading Test comprised the standardized measure used in the study.

Readability of the passages from the reader and the Gray Oral Reading Test was established using a computer program developed by Education Service Unit #3 in Millard, Nebraska. Readability formulas calculate grade equivalents using the average number of words in a sentence, average number of syllables per word and number of words with six or more letters. The probes from the reader were .2 to .5 grade level above the Gray Oral Reading Test passages.

Procedure

Subjects were all seen during the month of March, 1988. All

subjects read to the examiner on an individual basis. Subjects were given standardized instructions (see Appendix A). Passages from the reader and the Gray were alternated and the order was randomly varied.

Each child was allowed to read CBA probes for one minute and the number of words read was recorded. On Forms A, B, and C of the Gray, each subject read the entire passage and the time it took to complete the passage was recorded. Errors in reading on all six passages were recorded according to the method outlined in the manual of the Gray. Total number of errors rather than types of errors were recorded.

The first 20 children were tape recorded so interrater and intrarater reliability could be established. Interrater reliability was established by having a third grade teacher familiar with the Gray error recording procedures listen to the tapes and record the data. Results were then compared to the original data and a second set of data recorded by the examiner. Intrarater reliability was determined by comparing the original data for the first 20 children with the data obtained when the examiner listened to the tape recordings ten days after the recordings were made. Reliabilities were established following the first 20 subjects before continuing the study. Intrarater reliability was high with correlations above .97 for number of errors recorded and the speed measures for both the curriculum-based probes and the Gray Interrater reliability correlations between the experimenter and the teacher were above

.95 on all measures.

In accordance with curriculum-based procedures, the median score from the three passages from the reader and the median score from the three passages from the Gray were used for correlation with the teacher ratings in order to test the hypothesis.

Each child's reading teacher was asked to rate the child's speed, accuracy and overall reading skills on a seven-point scale (See Appendix B).

Speed was rated from one to seven with a rating of one reflecting a reading rate slower than 95% of the class and seven faster than 95% of the class. Accuracy was rated according to the number of errors the teacher perceived the student made per minute of oral reading with one indicating eight or more errors and seven indicating a student who rarely or never makes errors. The seven point scale for overall reading skills rated a students' reading skills according to the grade received in class with one being very poor or failing work and seven superior or A+ work.

Results

Statistical analysis was completed using the Dyna-Stat computer program published by Dynamic Microsystems, Inc.

The mean, standard deviation and variance of the subjects' performance on the CBA reading probes and the Gray are shown in Table 1.

The subjects' mean number of words read per minute on the passages from the third grade reader was 117.9 with a standard deviation of 36.8. This is well above the mastery level cut off of

50 words read per minute recommended by Deno and Mirkin (1977) for third graders. It also falls into the mastery level of over 100 words read per minute recommended by the same authors for grades four and above. This high average rate indicates that the average third grade child's speed of reading was well above an instructional level in the Glenwood Schools.

Table 1

Means, Standard Deviations, and Variance of Subjects' Performances on the Curriculum-Based Reading Probes and the Gray Oral Reading Test Passages

	Mean	Standard Deviation	Variance
Curriculum-Based Reading Probes			
Speed (Number of Words read per minute)	117.9	36.8	1353.6
Accuracy (number of errors made)	4.9	2.8	7.8
Gray Oral Reading Test Passages			
Speed (number of seconds to read passage)	35.6	17.0	290.2
Accuracy (number of errors made)	3.6	3.2	10.7

Accuracy was not at the mastery level of two errors recommended by Deno and Mirkin (1977). The mean number of errors of 4.9 with a standard deviation of 2.8 fell in the instructional level which allows three to seven errors.

On the Gray subjects read the 67-72 word passages in an average of 35.6 seconds with a standard deviation of 17.0 seconds. Assuming the subjects had continued reading at the same rate of speed this would translate into a rate of 125 words per minute which is comparable to the mean rate obtained from the CBA probes. A correlation of -.848 exists between number of words read on the curriculum probes and number of seconds required to read the Gray passages. Since the Gray was not administered as a test, no grade equivalent for the subjects was calculated. The mean errors made on the Gray passages was 3.6 with a standard deviation of 3.2 errors.

Table 2

Means, Standard Deviations, and Variance of Teachers' Predictions on a Seven Point Scale of Speed, Accuracy, and Overall Reading Skills

	Mean	Standard Deviation	Variance
Teacher Ratings of			
Speed	4.8	1.6	2.7
Accuracy	5.2	1.4	1.9
Overall Reading Skills	4.8	1.3	1.6

The descriptive statistics for the teachers' ratings of speed, accuracy and overall reading skills are indicated in Table 2. The mean ratings point out that the teachers rated all three skills above average which corresponds to the students' above

average speed performance. Mean ratings of speed and overall reading skills were exactly the same. Teachers tended to minimize predictions of errors, predicting an average of 3 errors per minute of oral reading which compares to an actual average of 4.9 errors made per minute when students read selections from their basal reader.

Table 3

Correlations Between Teacher Ratings of Speed, Accuracy, and Overall Reading Skills with Student Reading Performance on the Curriculum-Based Probes and the Gray Oral Reading Test

		Curriculum-Based		Gray Passages	
		# Words Read	# Errors	Seconds	# Errors
Teacher Ratings	Speed	.627	-.252	-.626	-.512
	Accuracy	.506	-.452	-.579	-.596
	Overall	.652	-.402	-.629	-.616

Table 3 reflects the correlation matrix of the teacher ratings and the students' performances on the curriculum probes and the Gray. Results indicate significant correlations between teacher rating of speed and student reading speeds. It is appropriate to look at the absolute value of the correlations since two measures of speed were used, i.e., number of words read per minute on the CBA reading probes and number of seconds required to read the Gray passages. The correlation between the teachers' speed ratings and

the performance on CBA is positive, while the correlation between the teachers' rating of speed and the performance on the Gray is negative. When the correlations are looked at as absolute values, it can be determined that the teachers' rating of speed equally predicted students' actual speed of reading on both the probes from the basal reader (r=.627)and the passages from the Gray (r=.626). Teacher ratings of speed correlated at a high level with errors on the Gray (r=.512), but unexpectedly at a much lower level with errors on the CBA passages (r=.252).

When asked to specifically rate accuracy, teachers' ratings correlated better with the actual errors made on the Gray than on the CBA probes. Teachers' ratings of accuracy correlated -.596 with number of errors made on the Gray passages and -.452 with the CBA probes. A negative correlation was expected since a high rating on the teachers seven point scale reflects a prediction of fewer errors.

Teachers' ratings of overall reading skills correlated with the CBA probes and Gray passages .652 and -.629, respectively. Ratings of overall reading skills correlated -.616 with errors made on the Gray and -.402 with errors made on the curriculum-based probes.

Table 4 indicates the correlations of the performance on the CBA probes with the performance on the Gray passages. Number of words read in one minute on the curriculum-based probes correlated highly (-.848) with the number of seconds required to read the Gray passages. The longer it took students to read the Gray, the more

errors made (r=.778). The correlation between speed and errors was not nearly as strong on the curriculum-based probes as on the Gray. The relationship was negative indicating the more words read the fewer errors made with a correlation of -.335.

Interestingly, the number of words read on the curriculum-based probes correlated higher with the errors made on the Gray passages (r= -.646) than with the errors made on the curriculum-based probe (r= -.335). Therefore the faster a student read the curriculum probes the fewer errors they made when reading the Gray passages.

Table 4

Correlations Between Performances on the Curriculum-Based Reading Probes and Performance on the Gray Oral Reading Test Passages

		CURRICULUM		GRAY	
		# Words Read	Errors	Seconds	Errors
Curriculum	# Words Read		-.335	-.848	-.646
	Errors	-.335		.342	.571
Gray	Seconds	-.848	.342		.778
	Errors	-.646	.571	.778	

A backward stepwise regression yielded multiple r's of the independent variables (teacher ratings) for each dependent variable: 1) seconds required to read the Gray passage, 2) number of errors made on the Gray, 3) words read per minute in the curriculum-based probe, 4) errors made in the curricula. The

multiple r's were compared using a t-test. Results indicate that teacher ratings equally predicted the number of seconds it took students to read the Gray passages, the errors made on the Gray, and the words read per minute in the curriculum-based probes. There was a significant difference between the multiple r's of these three dependent variables and the multiple r of the errors made in the curriculum-based reading probes. The difference reflects the apparent difficulty teachers had in accurately predicting the number of errors students make when reading from their classroom curriculum.

Table 5

A Comparison of Multiple R's of Teachers Ratings of Speed, Accuracy and Overall Reading Skills and Performance Measures for Curriculum-Based Reading Probes and Passages from the Gray Oral Reading Test

Dependent Variables Compared	t
Words/minute (CB) - Seconds (Gray)	0.4062 (not significant)
Words/minute (CB) - Errors (Gray)	1.367 (not significant)
Words/minute (CB) - Errors (CB)	6.467 ($p<.001$)
Seconds (Gray) - Errors (Gray)	0.875 (not significant)
Seconds (Gray) - Errors (CB)	5.66 ($p<.001$)
Errors (Gray) -Errors (CB)	4.5 ($p<.001$)

The beta weights of each teacher rating were calculated in order to have a standardized coefficient for comparison of which ratings best predicted each dependent variable. Table 6 lists the beta weights of each teacher rating for each dependent variable.

In general, the teachers' ratings of overall reading skills were the best predictors of performance on the CBA probes, while

ratings of speed and overall reading skills equally predicted speed on the Gray and ratings of accuracy best predicted errors made on the Gray.

Table 6

Beta Weights of Teacher Ratings for Dependent Variables

	Beta Weight		
Dependent Variable	Speed	Accuracy	Overall
Errors made on Gray	.289	-.460	-.269
Errors made on C-B	*	-.303	-.386
Seconds/Gray	-.308	-.118	-.308
Words per Minute/C-B	.358	-.115	.470

*too insignificant to be calculated

Discussion

Results of the study did not support the original hypothesis that teachers' ratings of speed and accuracy in oral reading would better predict the students' performance on reading passages drawn from the basal reader than their performance on a typical standardized reading test like the Gray. Teachers' ratings of overall reading skills best predicted performance on CBA probes. Ratings of speed correlated well with speed data on both the Gray passages and the CBA probes. Teachers had significantly more difficulty predicting the number of errors made when reading from the classroom curriculum than in predicting speed of reading or errors made on the standardized oral reading test.

Results suggest content validity was not a significant factor

affecting teacher's ability to predict performance. Teacher ratings of overall reading skills, not ratings of speed and accuracy, best predicted CBA performance. Jensen (1984, 1985) would argue that even though teachers attempted to predict specific skills, they actually rated global abilities, i.e., general intelligence. Therefore, the tests' content validity may play a smaller role than originally thought in assessing a test's accuracy of measurement.

The failure of the present field study to support the advantage of CBA content validity over a standardized test of reading causes critical examination of the current uses of CBA. Teachers in this field study were significantly poorer in predicting errors on the CBA probes than on the Gray passages. This result was unexpected and indicates the possibilty that teachers assume fast readers are accurate readers. Teachers may also interpret errors differently. Errors are not easily defined and are therefore difficult to monitor. It follows that if teachers have trouble determining the number of errors a student makes, they may not be cognizant of the types of errors made.

Since teachers have difficulty predicting errors it is important for some type of error analysis to be included in an assessment. There is no CBA model that successfully integrates error analysis into the evaluation process. The Curriculum-Based Evaluation model (Howell, 1986; Howell & Morehead, 1987) was developed to address errors made by students. Unfortunately the model does not outline any procedures for norming, therefore, it

cannot be determined if errors made by an individual differ significantly from peers' performances.

CBA alone does not constitute a comprehensive evaluation. CBA is a method of providing information on which educational decisions can partly be based. CBA uses fluency as the focus for decision-making (Deno, et. al., 1982; Idol, Nevin, & Paolucci-Whitcomb, 1986). The only descison that can be made from a fluency measure is whether or not a student's speed of reading is significantly slower than peers. When fluency is determined to be a problem, then hypotheses must be generated concerning why. Information from other sources must be available to draw from in forming hypotheses. If CBA is used in conjunction with standardized tests and the evaluation includes an analysis of teaching techniques and the instructional environment (Lentz & Shapiro, 1986; Shapiro & Lentz, 1985) the danger of misuse is eliminated and remediation recommendations can be made.

Use of CBA brings up the expectations that local norms will be developed, but in reality very few schools have established local norms. Pine County, Minnesota has been a forerunner in the use of CBA (Germann & Tindal, 1985) and is one of the few places in the United States with extensive local norms. Presently schools in Oregon and Iowa have begun developing local norms, but sites are scattered. Lack of local norms is probably one reason why CBA is not widely used. Wesson, King & Deno (1984) surveyed learning disabilities teachers and found although 80% were aware of CBA, only about half were using it.

The importance of establishing local norms when using CBA was underscored by the results of this field study. There was a large discrepancy between the subjects' mean reading speed of 117.9 words per minute and the instructional level recommended by Deno and Mirkin (1977) for third-graders of 50 words per minute. In the subject population a student two standard deviations below the mean speed would be close to mastery speed using Deno and Mirkin's guidelines. Clearly, peer referencing is needed to provide an accurate interpretation of an individual's skills.

Finally, CBA makes the assumption that the classroom curriculum is the most appropriate. The quality of the curriculum is not questioned and because the curriculum is used in the evaluation, there is no way to objectively assess it. Generalization of skills and learning styles are not taken into consideration. If we truly want to meet the needs of the individual student and maximize learning, then we must be wary of molding a student to fit a curriculum instead of molding the curriculum to fit the student.

Despite its apparent disadvantages, CBA does have a place in comprehensive evaluation. Through frequent progress monitoring CBA can provide feedback on the effectiveness of intervention procedures. Progress monitoring is also sensitive to small increments of progress which is often difficult for a teacher to determine subjectively. A student's rate of progress can be graphed and compared to baseline or with peers if local norms are developed. CBA is not a magical method of assessment, but it is practical and provides relevant information for educational decision making.

References

Angoff, W.H. (1971). Scales, norms and equivalent scores. In R.L. Throndike (Ed.), Educational Measurement (pp. 508-600). Washington, D.C.: American Council on Education.

Blankenship, C.S. (1985). Using curriculum-based assessment data to make instructional decisions. Exceptional Children, 52, 233-238.

Colardarci, T. (1986). Accuracy of teacher judgement of student responses to standardized test items. Journal of Educational Psychology, 78, 141-146.

Deno, S.L. (1985). Curriculum-based measurement: The emerging alternative. Exceptional Children, 52(3), 219-232.

Deno, S.L. (1986). Evaluation of individual student programs: A new role for school psychologists. School Psychology Review, 15(3), 358-374.

Deno, S.L. & Mirkin, P.K. (1977). Data-based program modification: A manual. Reston, VA: Council for Exceptional Children.

Deno, S.L., Mirkin, P.K., & Chiang, B. (1982). Identifying valid measures of reading. Exceptional Children, 49(1), 36-45.

Elliott, S.N. & Bretzing, B.H. (1980). Using and updating local norms. Psychology in the Schools, 17, 196-201.

Fuchs, L. & Deno, S. (1981). The relationship between curriculum based mastery measures and standardized achievement tests in reading. (Research Report No. 57) Mineapolis, Minnesota: University of Minnesota, Institute for Research on Learning Disabilities.

Fuchs, L.S., 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., & Deno, S.L (1982) Reliability and validity of curriculum-based informal reading inventories. Reading Research Quarterly, 18, 6-26.

Germann, G. & Tindal, G. (1985). An application of curriculum-based assessment: The use of direct and repeated measurement. Exceptional Children, 52(3), 244-265.

Gickling, E.E. & Thompson, V.P. (1985). A personal view of curriculum-based assessment. Exceptional Children, 52(3), 205-218

Howell, K.W. (1986). Direct assessment of academic performance. School Psychology Review, 15, 324-335.

Howell, K.W., & Morehead, M.K. (1987). Curriculum-based evaluation for special and remedial education. Columbus, OH: Charles Merrill.

Idol, L., Nevin, A., & Paolucci-Whitcomb, P. (1986). Models of curriculum-based assessment. Rockville, MD: Aspen Publishers.

Idol-Maestas, L., Lloyd, S., & Lilly, S. (1981). A non-categorical approach to direct service and teacher education. Exceptional Children, 48, 213-220.

Iowa Learning Disabilities Consortium (1989). Criterian validity of curriculum based assessment and correlation with teacher ratings and Iowa Test of Basic Skills, Des Moines, IA: Iowa Department of Education.

Jenkins, J. & Pany, D. (1978). Standardized achievement test: How useful for special education? Exceptional Children, 44, 448-453.

Jensen, A. R. (1984), Test Validity: g versus the specificity doctrine. Journal of Social and Biological Structures, 7, 93-118.

Jensen, A. R. (1985), Compensatory Education and the theory of intellignece. Phi Delta Kappan, April, 554-558.

Kamphaus, R. W. & Lozano, R. (1984). Developing local norms for individually administered tests. School Psychology Review, 13, 491-498.

Leinhardt, G. & Seewald, A.M. (1981). Overlap: What's tested, what's taught? Journal of Educational Measurement, 18, 85-95.

Lentz, F.E., Jr. & Shapiro, E.S. (1986). Functional assessment of the academic environment. School Psychology Review, 15(3), 346-357.

Lovitt, T.C. & Fantasia, K. (1980). Two approaches to reading program evaluation: A standardized test and direct assessment. Learning Disability Quarterly, 3, 77-87.

Marston, D. & Magnusson, D. (1985). Implementing curriculum-based measurement in special and regular education settings. Exceptional Children, 52(3), 266-276.

Marston, D., Mirkin, P., & Deno, S. (1984). Curriculum based measurement: An alternative to traditional screening, referral, and identification. The Journal of Special Education, 18(2), 109-117.

Neville, L., & Hamm, N.H. (1985). Content validity of the WRAT and grade level achievement of primary grade students. Journal of School Psychology, 23, 91-93.

Shapley, C.F., & Edgar, E. (1986). Teachers' ratings vs. standardized tests: a emperical investigation of agreement between two indices of achievement. Psychology in the Schools, 23, 106-111.

Shapiro, E. & Lentz, E. (1987, October). Relevand educational assessment and intervention model. Presented at Project Re-Aim. Workshop conducted for school psychologists and educational consultants in Area Education Agencies 13 and 14, Atlantic, Iowa.

Shapiro, E.S. & Lentz, F. E., Jr. (1985). Assessing academic behavior: A behavioral approach. School Psychology Review, 14, 325-338.

Shinn, M. (1988). Development of curriculum-based local norms for use in special education decision-making. School Psychology Review.

Webster, R.E., McInnis, E.D. & Craver, L. (1986). Curriculum biasing effects in standardized and criterion - reference reading achievement tests. Psychology in the Schools, 23, 205-213.

Wesson, C.L., King, R.P., & Deno, S.l. (1984). Direct and frequent measurement of student performance: If it's good for us, why don't we do it? Learning Disability Quarterly, 7, 45-48.

Appendix A

Instructions Administered to Subjects

"When I say 'start,' begin reading aloud at the top of this page. Read across the page (DEMONSTRATE BY POINTING). Try to read each word. If you come to a word you don't know, I'll tell it to you. Be sure to do your best reading. Are there any questions?"

Appendix B

Teachers' Rating Scales for Speed, Accuracy and Overall Reading Skills

1	2	3	4	5	6	7
slower than 95% of class	slower than 90% of class	slower than 70 of class	50%	faster than 70% of class	faster than 90% of class	faster than 95% of class

1	2	3	4	5	6	7
8 or more errors per min. of oral reading	7 errors per min. of oral reading	6 errors per min. of oral reading	4-5 errors per min. of oral reading	3 errors per min. of oral reading	2 or fewer errors per min. of oral reading	rarely or never makes errors

1	2	3	4	5	6	7
very poor, failing work	Below Average (D work)	Low Average (D+ or C-)	Average (C work)	Above Average (B work)	Very Good (A or A-work)	Superior (A+ work)

Appendix C

Curriculum-Based Probes

One of the new lambs was very weak, so Mary took the lamb into their house to care for it. First, she wrapped the little animal in some soft cloths. Then she sat down close to the fire with the lamb in her lap. All day and all night, Mary watched the lamb. Several times she tried to give it some tea or warm milk.

At first the lamb was too weak to drink, but toward evening, it took its first swallow of warm tea. During the night, it was able to drink a little more. By morning, the lamb was stronger. It was able to stand up for the first time. It was really hungry, too, and began to drink lots of milk.

"The lamb will be all right now," Mary thought. "There's no reason to worry anymore."

From that day on, the little lamb grew steadily stronger. Whenever Mary called, it came trotting to her side. It followed her around and lay down at her feet when she sat. The lamb wasn't just one of the farm animals. It had all the freedom of any well-loved pet.

One day Mary and her brother decided to take the lamb to school. Off they went with the lamb trotting along behind them. They reached the school ahead of their teacher and most of their schoolmates. Only a few friends saw them take the lamb into the schoolhouse.

Mary hid the lamb under her desk and covered it so that its white coat—called "fleece"—did not show.

Appendix C, continued

Exciting things often happen in circuses, and the Los Muchachos circus is no different. The boys of this circus have learned something important that every good performer knows. Hard work can make exciting things happen. The audience, watching, may think, "How brave that performer is! He makes it look so easy!"

The performer knows that he does well mostly because he has had good training. Suppose you were a boy like José Cabreas, with a high-wire act. Your training would teach you the best way to use your feet and toes to safely cross the wire strung high above the ground. You would learn the best way to breathe and to use the pole for balance. Many other skills would be a part of your training.

Then you would practice your act—and practice—and *practice!*

Do the boys in the circus sometimes become shaky about doing their acts? Are there days when they'd rather not perform at all? Maybe, but they go on anyway. If you were in the audience, you could not tell for sure how they felt inside. Indeed, as you watched them, you might be more upset than they!

Appendix C, continued

For more than an hour the boat slid gently through the waters of the river. After a while, however, the water began to move faster. The little boat picked up speed. From time to time sharp rocks appeared in the river, and Mai Sun and Mr. Bin had to paddle carefully around them.

"Rainbow Falls are coming up," said Mr. Bin. "Soon we will have to take the boat out of the water and carry it around the falls. If we went over the falls in this boat, we'd be smashed to pieces."

The closer the boat drifted toward the falls, the meaner the river became. The water swirled and churned. It slapped against the sharp rocks and filled the air with a glittering mist. Then, all at once, Su Ling heard a thundering roar. It was Rainbow Falls.

"Head for the shore," shouted Mai Sun. "The falls are not far away!"

As they turned for the shore, the boat brushed against a rock. The rock lifted one side of the boat out of the water. When the boat came down, Su Ling heard Mr. Bin shout. "My paddle!" he cried. "I've lost my paddle!"

Su Ling could see the paddle bobbing near the boat. Mr. Bin tried to grab for it, but it stayed just out of his reach.

"Now we're in real trouble," Mr. Bin cried. "With only one paddle, we'll never get to shore."