DOCUMENT RESUME

ED 225 113 CS 006 954

AUTHOR Tindal, Gerald; And Others
TITLE Curriculum Differences in Direct Repeated Measures of Reading.
INSTITUTION Minnesota Univ., Minneapolis. Inst. for Research on Learning Disabilities.
SPONS AGENCY Office of Special Education and Rehabilitative Services (ED), Washington, DC.
REPORT NO IRLD-RR-93
PUB DATE Oct 82
CONTRACT 300-80-0622
NOTE 39p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Basal Reading; Curriculum Evaluation; Elementary Education; Measurement Techniques; Oral Reading; *Reading Achievement; *Reading Materials; *Reading Research; *Reading Tests; *Textbook Research; Word Lists

ABSTRACT

Using 660 elementary school students from six school districts as subjects, a study examined student performance in reading aloud from four different reading curricula (Houghton Mifflin; Ginn 720; Holt, Rinehart & Winston; and Scott, Foresman), and the sensitivity of scores within each curriculum to growth across grade levels. Three different reading measurement tasks were developed. Two of the tasks were curriculum based; the other, which served as a common measure, was not curriculum bound. The two basal reading measurement tasks consisted of a reading passage and a vocabulary list; the common measure was a word list. The testing of the students was conducted within the first month of school. Results indicated that in all grades there appeared to be differences among the reading series in at least one of the measurement tasks--reading word lists or reading passages. The most consistent finding was that, when differences existed, Scott, Foresman and Houghton Mifflin appeared to be in the low student performance set, while Ginn 720 and Holt, Rinehart & Winston appeared in the high student performance set. This was especially true when the task utilized a word list rather than a reading passage. Results further indicated that the tasks for all four curricula adequately reflected student growth across grade levels. (HOD)

Research Report No. 93

CURRICULUM DIFFERENCES IN DIRECT REPEATED MEASURES OF READING

Gerald Tindal, Doug Marston, and Stanley L. Deno

Institute for Research on Learning Disabilities

University of Minnesota

Gary Germann

Pine County Special Education Cooperative

October, 1982

Abstract

The performances of 660 elementary students in six school districts on two curriculum-based reading aloud tasks and one non-curriculum-based (common) measure were examined. Four different reading curricula were compared. In addition, the sensitivity of the reading measures to growth across grade levels was examined. Results indicated that differences existed in students' performance on the curriculum-based reading tasks at all grades. These tasks for all four curricula adequately reflected student growth across grades. The need for a common, non-curriculum-based, measure for comparisons among students using different reading curricula is emphasized.

Director: James E. Ysseldyke

The Institute for Research on Learning Disabilities is supported by a contract (300-80-0622) with the Office of Special Education, Department of Education, through Title VI-G of Public Law 91-230. Institute investigators are conducting research on the assessment/decision-making/intervention process as it relates to learning disabled students.

During 1980-1983, Institute research focuses on four major areas:

- Referral
- Identification/Classification
- Intervention Planning and Progress Evaluation
- Outcome Evaluation

Additional information on the Institute's research objectives and activities may be obtained by writing to the Editor at the Institute (see Publications list for address).

The research reported herein was conducted under government sponsorship. Contractors are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent the official position of the Office of Special Education.

Curriculum Differences in Direct Repeated Measures of Reading

Concern about the usefulness of standardized achievement tests frequently is expressed by educators. Specifically, the notion that standardized achievement tests aid school personnel in making accurate placement, instructional planning, and progress evaluation decisions has been challenged (Salvia & Ysseldyke, 1981). In the area of reading, Chall (1970) noted that many standardized tests distort our view of the student's real level of reading. Too much emphasis is placed upon the norm-referenced scores derived from such tests, scores that provide little or no information about actual reading skills. Similarly, Farr (1969) indicated that the most serious deficiency of standardized reading tests is their lack of diagnostic validity.

These warnings have been substantiated by research findings. Eaton and Lovitt (1972) found that student achievement in reading is frequently a function of the test used for assessment. Tallmadge (1977) demonstrated that low-achieving students, although appearing to make adequate progress on standardized reading tests, actually fall farther behind their peers. Hively and Reynolds (1975) stated that tests are not adequately designed to measure growth. A review of factor analytic studies focusing on standardized group reading tests convinced Spache (1976) that the tests measured few specific reading skills. He stated that "diagnostic profiles based on comprehension subskill performance often are relatively meaningless" (p. 252). Finally, Jenkins and Pany (1978) examined the overlap between the content of several standardized achievement tests of reading and commercial reading series. Upon finding extreme variability in the content of the tests, the authors concluded that because of the

presence of curriculum bias, reading achievement tests yield unsatisfactory assessment information:

> What educators need is an instrument to measure learning that is sensitive to curricular differences. Some form of criterion referenced or curriculum based assessment may provide the solution. Frequent and direct measures of a child's performance in a specific curriculum should reveal what skills within the curriculum have or have not been mastered, as well as provide some index of progress that would be sensitive to what was being taught. (p. 453)

The problematic nature of standardized achievement tests has indeed prompted educators to search for assessment procedures sensitive to both curriculum differences and student improvement. Frequent and direct measurement in a child's curriculum possesses these desired characteristics (Becker & Englemann, 1976; Deno & Mirkin, 1977; White & Haring, 1981).

In the measurement of reading achievement, advocates of frequent and direct measurement suggest that the skill area assessed should be fluency in reading aloud from either word lists or passages (Jenkins, Deno, & Mirkin, 1979). There is substantial evidence at the elementary grade levels to support this recommendation. Guthrie and Tyler (1978) found a substantial correlation between reading rate and comprehension. Deno, Mirkin, and Chiang (1982) corroborated these results and, in addition, found performance in reading aloud from both text and word lists to be sensitive to student improvement across six elementary grade levels.

Despite these useful claims of technical adequacy for reading rate measures, there is a need for further research before such practices are implemented. An unanswered question is the issue of whether reading aloud scores vary with the curriculum as do

standardized achievement test scores, thus revealing important differences in the curriculum. The issue becomes significant in those situations where students performing in different curricula are compared to one another. If student performance varies as a function of the curriculum, what measures can be used for cross-curricula comparisons?

A second unresolved issue is sensitivity to growth. Deno, Mirkin, and Chiang (1982) demonstrated the discriminant validity of correct performance in reading aloud between elementary grade levels. The research, however, was conducted on only a few selected text passages. No effort was made to determine whether the same results could be obtained consistently with a variety of commonly used curricula.

The purpose of this research was to undertake an analysis of both issues. In this investigation we examined the comparability of student performance in reading aloud from four different reading curricula and the sensitivity of scores within each curriculum to growth across grade levels.

Method

Subjects

A random sample of 660 elementary students attending six different school districts in a rural midwestern education cooperative served as subjects for this study. Class rosters were obtained for each grade (1-6) in the six districts and students' names numbered. Using a table of random numbers, equal numbers of students were selected from all classes for each grade. In four districts, 20

students from each grade were selected; 15 students were selected from each grade in the remaining two districts. No attempt was made to obtain equal representation of males and females.

Materials

Three different reading measurement tasks were developed. Two of the tasks were curriculum-based; the other, which served as a common measure, was not curriculum bound. The two basal reading measurement tasks consisted of a reading passage and a vocabulary word list; the common measure was a word list.

Four different basal reading series were in use in the six school districts, with two districts using the same reading series as two other districts. The four different series included: Houghton-Mifflin (1971 edition; Durr, Lapere, & Brown, 1971), Ginn 720 (1979 edition, Clymer & Fenn, 1979), Holt-Rinehart-Winston (1980 edition; Weiss, Everetts, Cruickshank, & Steuer, 1980), and Scott-Foresman (1981 edition; Aaron, Jackson, Riggs, Smith, & Tierney, 1981). Using the publishers' criteria for dividing the various books and magazines into appropriate grade levels, material was selected that corresponded to the beginning of each grade level.

Reading passages. The reading passage measure was developed by randomly selecting enough pages from this material (magazine, book, level) to present a minimum of 50 words in grade 1, 100 words in grade 2, and 200 words in grades 3-6. The only criterion for screening the pages selected was that no special dialogue or poetry be present in the passage. Two copies of each passage were made, one from which the student read and one on which the tester followed along. The tester's

copy included a cumulative count of the number of words in each successive line.

Word lists. Two types of word lists were developed. The basal word list measure consisted of a random selection of 150 words from the pool of all vocabulary words appearing within each of the grade levels from which text passages were drawn. The lists were cumulative and stratified, with 60% of the words sampled from grade level and 40% consisting of words sampled equally from previous grades. For example, a fifth grade list of 150 words would include 90 words from the fifth grade level and 15 words from each of the first through fourth grade levels. Again, two copies of each word list were made, one for the student to read, and a follow-along list for the tester.

The common word list consisted of a random selection of 150 words appearing in the core list of words in the Basic Reading Vocabulary (Harris & Jacobson, 1972), a computer based compilation of over 5000 words appearing in commonly used basal readers. The domain of words from which the list was constructed was limited to words from preprimer through grade four. Two copies of each list were made, one for the student to read, and a follow-along list for the tester.

Procedures

The testing of the 660 students was conducted within the first month of school. Ten educational aides were trained in the measurement procedures during a two-hour session and given all the necessary materials, including word lists, reading passages, recording forms, and stop watches. Each aide was assigned to a school, given a list of students, and a time schedule. All testing was done on an

individual basis and was completed within one week. Each session involved the administration of two reading passages, one basal word list, and one common measure. Within each testing session, the order of administration of the three measures was counterbalanced. Each of the reading measures was a one-minute sample. Directions given to the student were as follows:

> When I say start, begin reading at the top of the page. If you wait on a word for too long, I'll tell you the word. If you come to a word that you cannot read, just say pass and go on to the next word. Do not attempt to read as fast as you can. This is not a "speed reading" test. Read at a comfortable rate, and at the end of one minute I'll say "stop."

The testers followed along on their copy and marked errors of omission, substitution, and mispronunciation. If the student failed to continue within approximately 5 seconds, the tester prompted the student to skip that word and continue. Immediately following each one-minute measure, the student's performance was scored and recorded, with no feedback given to the student.

Results

Comparison of Districts Using the Same Curriculum

Prior to comparing the basal reading series with one another, the two pairs of districts using the same series were compared to determine whether any population differences existed. The two school districts using Ginn 720 and the two districts using Houghton-Mifflin were compared at each grade level on the two basal reading measures--word lists and passages. Findings from these analyses indicated that in three grades there was a significant difference in the performance of students in the two districts using the Houghton-Mifflin reading

series. In grade 2, differences between districts appeared on both passage and basal word list tasks in the Houghton-Mifflin reading series and on the Ginn 720 basal word list task. In grade 4, differences between districts appeared on Ginn 720 basal passage and word list and on the Houghton-Mifflin basal passage. Finally, in grade 5, there was a difference between the two districts on the Ginn 720 basal passage. When the two districts using Ginn 720 were compared on the common word list measurement task, significant differences appeared for grades 1 and 2.

Comparisons Among Curricula

Means and standard deviations for students on the three reading measurement tasks for each basal reading series appear in Table 1. One-way analyses of variance were conducted on these data to ascertain the extent to which any differences existed between the reading series. Significant results were tested using the Student-Newman-Keuls *a posteriori* test.

Insert Table 1 about here

For the reading passage data, there was a significant difference between the various readers in four of seven grade level comparisons (see Table 2). Only in grades 2 and 3 and averaging across all six grades was there no significant difference. The *a posteriori* tests revealed that in the four grades showing a difference, Scott-Foresman consistently appeared in the low set (characterized by lower student performance) while Ginn 720 and Holt-Rinehart appeared primarily in

the high set (characterized by higher student performance). Analysis of performance on the basal vocabulary word list produced many of the same findings (see Table 3), with a total of three grades and the average across grades showing a significant different among the reading series. In grades 1-3, no significant difference among the four series was indicated. The a posteriori tests again showed Scott-Foresman and Houghton-Mifflin in the low set, while Ginn 720 and Holt-Rinehart quite consistently appeared in the high set. Student performance on the common word list provides a direct measure of comparability of students' performance unconfounded by the effect of the basal reader. No significant differences occurred in any of the grade levels (see Table 4).

Insert Tables 2-4 about here

An analysis of covariance was conducted to adjust for any differences between populations and thus provide a more valid comparison of the different reading series. The results of this analysis appear in Table 5, along with the rank order of each reading series. The rank order was obtained following adjustment for performance on the common measure. As can be seen from the F-ratios and F-probabilities in Table 5, there were significant differences among the four basal readers, both when the dependent measure was reading from a passage and when it was reading a vocabulary word list. In all six grade levels and in the average across grades, there was a significant difference among the reading series for at least one of

the measures. In grades 2 and 5, using the basal reading passage, and in grade 1, on the basal word list, no significant differences appeared.

Insert Table 5 about here

Table 6 is a summary of the rank order data from Table 5. As is evident in Table 6, the four reading series tended to cluster quite consistently in a dichotomous fashion. For the basal reading passage, Scott-Foresman received mostly ranks of 4, while Ginn 720 and Holt-Rinehart primarily received ranks of 1 and 2. Houghton-Mifflin fell in the middle. On the basal word list, Scott-Foresman and Houghton-Mifflin consistently received lower ranks (3 and 4) while Ginn 720 and Holt-Rinehart received ranks of 1 and 2.

Insert Table 6 about here

The data presented in Tables 7 and 8 help clarify whether reading rate measures are sensitive to growth within all reading curricula. Our major premise was that reading rate should increase with grade level if the measure is sensitive to a child's progress. The grade level means for the four reading curricula appeared to increase with grade. To test this hypothesis, the grade level means for each measure within the curriculum were first subjected to an analysis of variance. All F-values reported in Table 7 were significant. A more crucial measure of sensitivity, however, is a test of linearity. If

reading rate is sensitive to progress, it should increase with grade. Reading rates on basal passages and basal word lists from Ginn 720, Scott-Foresman, Houghton-Mifflin, and Holt-Rinehart all displayed linear trends, significant at the .05 level (see Table 8). Reading rate within any of the four curricula appears to be sensitive to educational progress.

Insert Tables 7 and 8 about here

Discussion

The data from this study consistently indicated that there are differences in the four basal reading series when using reading rate as the dependent measure. In each of the various analyses, the results were similar. Using either a one-way analysis of variance or an analysis of covariance, differences were apparent among the reading series on both the reading passage and basal word list measurement tasks. For most grades, this difference was found on both basal reading measures.

In a few grades, a difference was found in one measure but not the other. There is no obvious reason why this incongruity appears. One possibility is that for those grades, the two basal reading tasks were differentially difficult as a result of the procedures used in generating the measurement materials. That is, by randomly selecting words and passages, provision is made for stimulus equivalence only over the long run. Any particular selection may be more or less difficult. Given the differences within any reading series in the

type of vocabulary (Carnine & Silbert, 1979) and the readability of successive passages (Fuchs, 1981), it is possible that measurement stimuli were generated that were not representative of the material in general. The result was that the passage selected was unusually easy or difficult and failed to differentiate students or corroborate previously found differences between the various series. For the vocabulary list, the frequency of regular or irregular words may have resulted in a task atypical of that series and level in general. Again, with the task changed in difficulty, there is less likelihood of differentiating between students and basal series in a manner similar to the general pattern of findings. For the reading passages, the great differences in the readability found by Fuchs (1981) may have been operating in this study also.

In general, this research revealed real differences in the reading series with no strong population sampling effect indicated. There did not appear to be any differences between students of various districts, using both the common measure and both basal reading measures. Because two of the districts in this study were using the same reading series as two other districts, it was possible to compare directly these districts on all three measures. If differences appeared on the basal reading measures, the evidence supporting any differences in reading series would be seriously weakened.

In two districts, all elementary grades were using Ginn 720, while two other districts were using Houghton-Mifflin. A comparison of performance in these four districts revealed very few differences. In some grades (4 and 5), a difference appeared in one of the measures

but not the others. In the use of any random selection procedures, equivalence is obtained only over repeated samplings and any given sample may be biased in one direction or the other. It is entirely possible that the two groups of students from these districts were not representative of each of their general populations or that the measurement stimuli generated were not equivalent in difficulty. As a result, more (or less) proficient students were selected and compared with each other on more or less comparable tasks. It is interesting to note that such an occurrence is actually more than likely given the number of districts and grades sampled. In all, 36 different populations (of students and grade level materials) were sampled. The likelihood of obtaining no population differences in all of these is low. However, the word lists and passages in these districts were not exactly the same, but rather sampled from within the same grade level book. Therefore, population differences were confounded with curriculum differences, which may explain the population differences that were found.

The findings obtained on the one-way analyses of variance were essentially corroborated when analyses of covariance were used. Using student performance on the common measure as a covariate, a more exact comparison of the reading series could be made without any influence of differential entry skills. In all grades, there appeared to be differences among the reading series, in at least one of the measurement tasks--reading word lists or reading passages.

This research attempted not only to determine whether differences existed between the reading series, but the ways in which the

differences were exhibited. The two analyses which examined this aspect included the Student-Newman-Keuls test and the rank ordering of basal series on the reading rates obtained. In the former (Tables 2-4), reading series are put into homogeneous subsets when there are no significant differences in the mean performance. With four basal readers being compared, the total number of possible subsets is four. However, this analysis revealed that when differences appeared among the series, the resulting grouping into sets was dichotomous, with one set characterized by lower student performance and the other characterized by higher student performance. The most consistent finding was that, when differences existed, Scott-Foresman and Houghton-Mifflin appeared in the low set, while Ginn 720 and Holt-Rinehart appeared in the high set. This was especially true when the task utilized a word list rather than a reading passage. Ginn 720 clearly fell within the high set on the word list task.

A similar grouping of series was done by rank ordering them in difficulty using the deviation of performance from the grade mean, after adjusting for differences on the covariate (Table 6). Again, the data showed a very consistent dichotomy. The four reading series fell into two groups--one group in which ranks of 1 and 2 consistently were obtained and one in which ranks of 3 and 4 consistently were obtained. Again, Scott-Foresman and Houghton-Mifflin showed the lowest performance and corresponding ranks, while Ginn 720 and Holt-Rinehart revealed higher performance and ranks of 1 and 2. On the basal passage, the majority of ranks were consistent with this dichotomy, while on the basal word list, all ranks dichotomize in like

manner. This finding coincides very well with the analysis by Carnine and Silbert (1979). They found that the number of irregular words introduced in the first 50 lessons of Ginn 720, Houghton-Mifflin (1974), and Scott-Foresman (1981) was 29, 38, and 125, respectively. It is possible that reading rates are highly related to the frequency of irregular words.

In conclusion, it appears that differences do exist between the four basal reading series; these differences indicate that Scott-Foresman (1981 edition) and Houghton-Mifflin (1974 edition) generate lower reading rates than those obtained from the use of Holt-Rinehart (1981 edition) or Ginn 720 (1979 edition). The implications of this finding are particularly noteworthy.

According to Jenkins and Pany (1978), there is differential overlap between various achievement tests and basal reading series. This finding, however, was a result of a content analysis of the tests and reading series and was predicated on certain assumptions concerning student learning and performance. To corroborate this analysis, empirical data on actual student performance were needed. The findings from the present research partially provide such evidence. Given that students perform differentially according to the reading series used, and that achievement tests sample differentially from these series, it is likely that such biases will be reflected in students' performance on the achievement measures. This is particularly true for tests emphasizing decoding and vocabulary words. In essence, achievement tests no longer can be considered as impartial, common measures on which comparisons can be made among

students working in different readers.

An alternative to this type of measurement system would be the use of curriculum-based assessment as advocated by Jenkins, Deno, and Mirkin (1979). As long as measurement and analysis of student performance remains within a reading series, there is no bias in comparisons among students. However, with the use of many different basal readers, some common measure is needed that does not differentially sample from any one basal reading series. In many educational cooperatives, in which the delivery of special education is in concert with several different districts, the use of this common measure is imperative for valid between-district comparisons. Indeed, one of the basic tenets of curriculum-based assessment is the use of local norms. The use of such a system will result in generation of valid data having direct instructional relevance. Not only should the norms be based on a comparable population, but also utilize a comparable set of stimulus materials for conducting assessment.

References

Aaron, I., Jackson, D., Riggs, C., Smith, R., & Tierney, R. The Scott-Foresman reading. Glenview, IL: Scott-Foresman, 1981.

Becker, W. C., & Englemann, S. Teaching 3: Evaluation of instruction. Chicago: Science Research Associates, 1976.

Carnine, D., & Silbert, J. Direct instruction reading. Columbus, OH: Charles E. Merrill, 1979.

Chall, J. Interpretation of the results of standardized reading tests. In R. Farr (Ed.), Measurement and evaluation of reading. New York: Harcourt, Brace & World, 1970.

Clymer, T., & Fenn, T. Reading 720 rainbow edition. Lexington, MA: Ginn, 1979.

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

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

Durr, W., Lapere, J., & Brown, R. Houghton-Mifflin readers. Boston: Houghton-Mifflin, 1971.

Eaton, M., & Lovitt, T. Achievement tests vs. direct and daily measurement. In G. Semb (Ed.), Behavior analysis and education. Lawrence, KS: University of Kansas, 1972.

Farr, R. Reading: What can be measured? Newark, DE: International Reading Association, 1969.

Fuchs, L. S. The concurrent validity of progress measures of basal reading material. Unpublished Ph.D. dissertation thesis. Minneapolis: University of Minnesota, 1981.

Guthrie, J. T., & Tyler, S. J. Cognition and instruction of poor readers. Journal of Reading Behavior, 1978, 10, 57-78.

Harris, A. J., & Jacobson, M. D. Basic elementary reading vocabularies. New York: MacMillan, 1972.

Hively, W., & Reynolds, M. C. (Eds.). Domain-referenced testing in special education. Reston, VA: Council for Exceptional Children, 1975.

Jenkins, J., & Pany, D. Standardized achievement tests: How useful for special education? Exceptional Children, 1978, 44(6), 448-453.

Jenkins, J. R., Deno, S. L., & Mirkin, P. Measuring pupil progress toward the least restrictive environment. Learning Disability Quarterly, 1979, 2, 81-91.

Salvia, J., & Ysseldyke, J. E. Assessment in special and remedial education. Boston: Houghton-Mifflin, 1981.

Spache, G. D. Diagnosing and correcting reading disabilities. Boston: Allyn & Bacon, 1976.

Tallmadge, G. K. Cautions to evaluators. In M. Wargo & D. Green (Eds.), Achievement testing of disadvantaged and minority students for educational program evaluation. New York: McGraw-Hill, 1977.

Weiss, B., Everetts, E., Cruickshank, S., & Steuer, L. The Holt basic reading system. New York: Holt, Rinehart, & Winston, 1980.

White, O. R., & Haring, N. G. Exceptional teaching (2nd ed.). Columbus, OH: Charles E. Merrill, 1980.

Table 1

Students' Performance on Various Reading Measurement Tasks Across Four Different Basal Reading Series

Measurement Task	Student's Grade Level 1 $\bar{X}$	1 SD	2 $\bar{X}$	2 SD	3 $\bar{X}$	3 SD	4 $\bar{X}$	4 SD	5 $\bar{X}$	5 SD	6 $\bar{X}$	6 SD	Average $\bar{X}$	Average SD
Reading Passage														
Houghton-Mifflin	13.9	14.8	49.9	37.7	72.4	43.1	99.1	31.6	126.3	34.1	123.2	39.5	80.8	52.9
Ginn 720	8.3	9.4	39.9	29.6	69.9	41.2	111.8	39.2	120.9	43.1	134.2	40.4	80.8	57.8
Holt-Rinehart	2.8	8.6	45.2	22.1	92.0	43.2	105.4	32.5	128.9	28.3	141.9	40.0	86.0	57.3
Scott-Foresman	5.7	5.3	49.4	38.2	81.6	41.3	83.1	36.4	99.5	38.5	110.0	33.1	71.5	48.6
Basal Word List														
Houghton-Mifflin	1.1	2.5	18.5	30.0	26.8	19.3	39.4	15.6	34.1	13.5	37.9	17.0	26.3	20.7
Ginn 720	1.4	2.4	19.3	19.6	32.0	20.7	44.8	22.0	48.1	22.3	48.1	23.8	32.3	26.1
Holt-Rinehart	1.2	2.5	21.3	18.8	35.4	16.5	45.9	11.9	50.5	16.8	50.8	14.5	34.2	22.9
Scott-Foresman	1.6	1.6	21.9	20.5	24.1	19.8	25.3	14.5	31.8	17.9	36.1	12.7	23.5	19.0
Common Word List														
Houghton-Mifflin	.6	3.0	13.1	18.0	31.4	21.9	53.2	19.4	65.5	15.9	64.0	18.5	37.9	30.2
Ginn 720	.6	1.1	10.7	14.3	29.0	21.6	45.7	21.1	61.8	23.7	73.4	23.9	36.8	32.5
Holt-Rinehart	.3	1.0	10.3	.86	35.1	19.5	50.0	14.8	67.5	18.5	71.3	16.7	39.1	30.6
Scott-Foresman	.1	.3	17.8	18.2	37.7	21.9	51.8	22.2	54.7	19.1	69.4	18.5	38.6	29.6

Table 2

One Way Analysis of Variance Comparing Basal Readers Using Performance on Reading Passage Measurement Task: F-Ratio, F-Probability, and Grouping into Homogeneous Subsets Using Student-Newman-Keuls

Grade Level	F-Ratio	F-Probability	Number of different sets	Name of Reading Series in Low Set	High Set
1	4.71	.00	2	SF, HR Ginn	HM
2	.68	.57	1		
3	1.21	.31	1		
4	3.02	.03	2	SF	Ginn
5	2.57	.05	2	SF	HM, Ginn HR
6	2.58	.05	2	SF, HM	Ginn, HR
Average	1.36	.245	1		

Table 3

One-Way Analysis of Variance Comparing Basal Readers Using Performance on Basal Word List Measurement Task: F-Ratio, F-Probability, and Grouping into Homogeneous Subsets Using Student-Newman-Keuls

Grade Level	F-Ratio	F-Probability	Number of different sets	Name of Reading Series in Low Set	Name of Reading Series in High Set
1	.29	.84	1		
2	.16	.93	1		
3	1.40	.25	1		
4	6.14	.00	2	SF	HR, Ginn HR
5	6.65	.00	2	HM, SF	Ginn, HR
6	3.54	.02	1	HM, SF	Ginn, HR
Average	6.50	.00	2	HM, SF	Ginn, HR

Table 4

One-Way Analysis of Variance Comparing Basal Readers Using Performance on Common Word List Measurement Task: F-Ratio, F-Probability, and Grouping into Homogeneous Subsets Using Student-Newman-Keuls

Grade Level	F-Ratio	F-Probability	Number of different sets	Name of Reading Series in Low Set	Name of Reading Series in High Set
1	.39	.76	1		
2	1.01	.39	1		
3	.84	.47	1		
4	.97	.41	1		
5	1.62	.19	1		
6	1.38	.25	1		
Average	.16	.93	1		

Table 5

Analysis of Covariance Comparing Basal Readers on Two Basal Reading Measurement Tasks Using Performance on a Common Reading Task as the Covariate

Grade Level	Basal Reading Passage - Ranking of Series*						Basal Word List - Ranking of Series*					
	F-Ratio	F-Probability	HM	Ginn	HR	SF	F-Ratio	F-Probability	HM	Ginn	HR	SF
1	4.58	.00	1	2	4	3	.38	.77	4	2	3	1
2	1.02	.39	1	3	2	4	2.84	.04	3	2	1	4
3	2.88	.04	3	2	1	4	29.04	.00	3	1	2	4
4	12.59	.00	3	1	2	4	17.20	.00	3	1	2	4
5	1.07	.37	1	2	3	4	20.62	.00	4	1	2	3
6	5.64	.00	2	3	1	4	6.63	.00	3	2	1	4
Average	9.75	.00	3	2	1	4	30.00	.00	3	2	1	4

*Ranks obtained following adjustment for performance on measure used as covariate.

Table 6

Comparison of Four Basal Reading Series on Number of Times

Represented by each Possible Ranking on Two Measurement Tasks

Reading Series	Basal Reading Passage Number of Times with Rank of*				Basal Word List Number of Times with Rank of*			
	1	2	3	4	1	2	3	4
Houghton-Mifflin	3	1	3	0	0	0	5	2
Ginn 720	1	4	2	0	3	4	0	0
Holt-Rinehart	3	2	1	1	3	3	1	0
Scott-Foresman	0	0	1	6	1	0	1	5

*Ranks obtained following adjustment for performance on measure used as covariate.

Table 7

Analysis of Variance Results for Grade Level Means Within Each Reading Series for Curriculum-Based Measures

Reading Series	Basal Passage F-value	Basal Passage Significance	Basal Word List F-value	Basal Word List Significance
Houghton-Mifflin	56.47	.001	29.17	.001
Ginn 720	52.41	.001	29.42	.001
Holt-Rinehart	42.63	.001	27.46	.001
Scott-Foresman	24.76	.001	11.37	.001

Table 8

Tests of Linearity for Grade Level Means Within

Each Reading Series for Curriculum-Based Measures

Reading Series	Basal Passage F-value	Basal Passage Significance	Basal Word List F-value	Basal Word List Significance
Houghton-Mifflin	3.68	.007	7.43	.001
Ginn 720	2.78	.029	2.82	.027
Holt-Rinehart	3.05	.021	3.96	.005
Scott-Foresman	3.57	.008	2.44	.051

PUBLICATIONS

Institute for Research on Learning Disabilities
University of Minnesota

The Institute is not funded for the distribution of its publications. Publications may be obtained for $3.00 per document, a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be prepaid.

Requests should be directed to: Editor, IRLD, 350 Elliott Hall; 75 East River Road, University of Minnesota, Minneapolis, MN 55455.

Ysseldyke, J. E. Assessing the learning disabled youngster: The state of the art (Research Report No. 1). November, 1977.

Ysseldyke, J. E., & Regan, R. R. Nondiscriminatory assessment and decision making (Monograph No. 7). February, 1979.

Foster, G., Algozzine, B., & Ysseldyke, J. Susceptibility to stereotypic bias (Research Report No. 3). March, 1979.

Algozzine, B. An analysis of the disturbingness and acceptability of behaviors as a function of diagnostic label (Research Report No. 4). March, 1979.

Algozzine, B., & McGraw, K. Diagnostic testing in mathematics: An extension of the PIAT? (Research Report No. 5). March, 1979.

Deno, S. L. A direct observation approach to measuring classroom behavior: Procedures and application (Research Report No. 6). April, 1979.

Ysseldyke, J. E., & Mirkin, P. K. Proceedings of the Minnesota round-table conference on assessment of learning disabled children (Monograph No. 8). April, 1979.

Somwaru, J. P. A new approach to the assessment of learning disabilities (Monograph No. 9). April, 1979.

Algozzine, B., Forgnone, C., Mercer, C. D., & Trifiletti, J. J. Toward defining discrepancies for specific learning disabilities: An analysis and alternatives (Research Report No. 7). June, 1979.

Algozzine, B. The disturbing child: A validation report (Research Report No. 8). June, 1979.

Note: Monographs No. 1 - 6 and Research Report No. 2 are not available for distribution. These documents were part of the Institute's 1979-1980 continuation proposal, and/or are out of print.

Ysseldyke, J. E., Algozzine, B., Regan, R., & Potter, M. Technical adequacy of tests used by professionals in simulated decision making (Research Report No. 9). July, 1979.

Jenkins, J. R., Deno, S. L., & Mirkin, P. K. Measuring pupil progress toward the least restrictive environment (Monograph No. 10). August, 1979.

Mirkin, P. K., & Deno, S. L. Formative evaluation in the classroom: An approach to improving instruction (Research Report No. 10). August, 1979.

Thurlow, M. L., & Ysseldyke, J. E. Current assessment and decision-making practices in model programs for the learning disabled (Research Report No. 11). August, 1979.

Deno, S. L., Chiang, B., Tindal, G., & Blackburn, M. Experimental analysis of program components: An approach to research in CSDC's (Research Report No. 12). August, 1979.

Ysseldyke, J. E., Algozzine, B., Shinn, M., & McGue, M. Similarities and differences between underachievers and students labeled learning disabled: Identical twins with different mothers (Research Report No. 13). September, 1979.

Ysseldyke, J., & Algozzine, R. Perspectives on assessment of learning disabled students (Monograph No. 11). October, 1979.

Poland, S. F., Ysseldyke, J. E., Thurlow, M. L., & Mirkin, P. K. Current assessment and decision-making practices in school settings as reported by directors of special education (Research Report No. 14). November, 1979.

McGue, M., Shinn, M., & Ysseldyke, J. Validity of the Woodcock-Johnson psycho-educational battery with learning disabled students (Research Report No. 15). November, 1979.

Deno, S., Mirkin, P., & Shinn, M. Behavioral perspectives on the assessment of learning disabled children (Monograph No. 12). November, 1979.

Sutherland, J. H., Algozzine, B., Ysseldyke, J. E., & Young, S. What can I say after I say LD? (Research Report No. 16). December, 1979.

Deno, S. L., & Mirkin, P. K. Data-based IEP development: An approach to substantive compliance (Monograph No. 13). December, 1979.

Ysseldyke, J., Algozzine, B., Regan, R., & McGue, M. The influence of test scores and naturally-occurring pupil characteristics on psycho-educational decision making with children (Research Report No. 17). December, 1979.

Algozzine, B., & Ysseldyke, J. E. Decision makers' prediction of students' academic difficulties as a function of referral information (Research Report No. 18). December, 1979.

Ysseldyke, J. E., & Algozzine, B. Diagnostic classification decisions as a function of referral information (Research Report No. 19). January, 1980.

Deno, S. L., Mirkin, P. K., Chiang, B., & Lowry, L. Relationships among simple measures of reading and performance on standardized achievement tests (Research Report No. 20). January, 1980.

Deno, S. L., Mirkin, P. K., Lowry, L., & Kuehnle, K. Relationships among simple measures of spelling and performance on standardized achievement tests (Research Report No. 21). January, 1980.

Deno, S. L., Mirkin, P. K., & Marston, D. Relationships among simple measures of written expression and performance on standardized achievement tests (Research Report No. 22). January, 1980.

Mirkin, P. K., Deno, S. L., Tindal, G., & Kuehnle, K. Formative evaluation: Continued development of data utilization systems (Research Report No. 23). January, 1980.

Deno, S. L., Mirkin, P. K., Robinson, S., & Evans, P. Relationships among classroom observations of social adjustment and sociometric rating scales (Research Report No. 24). January, 1980.

Thurlow, M. L., & Ysseldyke, J. E. Factors influential on the psychoeducational decisions reached by teams of educators (Research Report No. 25). February, 1980.

Ysseldyke, J. E., & Algozzine, B. Diagnostic decision making in individuals susceptible to biasing information presented in the referral case folder (Research Report No. 26). March, 1980.

Thurlow, M. L., & Greener, J. W. Preliminary evidence on information considered useful in instructional planning (Research Report No. 27). March, 1980.

Ysseldyke, J. E., Regan, R. R., & Schwartz, S. Z. The use of technically adequate tests in psychoeducational decision making (Research Report No. 28). April, 1980.

Richey, L., Potter, M., & Ysseldyke, J. Teachers' expectations for the siblings of learning disabled and non-learning disabled students: A pilot study (Research Report No. 29). May, 1980.

Thurlow, M. L., & Ysseldyke, J. E. Instructional planning: Information collected by school psychologists vs. information considered useful by teachers (Research Report No. 30). June, 1980.

Algozzine, B., Webber, J., Campbell, M., Moore, S., & Gilliam, J. Classroom decision making as a function of diagnostic labels and perceived competence (Research Report No. 31). June, 1980.

Ysseldyke, J. E., Algozzine, B., Regan, R. R., Potter, M., Richey, L., & Thurlow, M. L. Psychoeducational assessment and decision making: A computer-simulated investigation (Research Report No. 32). July, 1980.

Ysseldyke, J. E., Algozzine, B., Regan, R. R., Potter, M., & Richey, L. Psychoeducational assessment and decision making: Individual case studies (Research Report No. 33). July, 1980.

Ysseldyke, J. E., Algozzine, B., Regan, R., Potter, M., & Richey, L. Technical supplement for computer-simulated investigations of the psychoeducational assessment and decision-making process (Research Report No. 34). July, 1980.

Algozzine, B., Stevens, L., Costello, C., Beattie, J., & Schmid, R. Classroom perspectives of LD and other special education teachers (Research Report No. 35). July, 1980.

Algozzine, B., Siders, J., Siders, J., & Beattie, J. Using assessment information to plan reading instructional programs: Error analysis and word attack skills (Monograph No. 14). July, 1980.

Ysseldyke, J., Shinn, M., & Epps, S. A comparison of the WISC-R and the Woodcock-Johnson Tests of Cognitive Ability (Research Report No. 36). July, 1980.

Algozzine, B., & Ysseldyke, J. E. An analysis of difference score reliabilities on three measures with a sample of low achieving youngsters (Research Report No. 37). August, 1980.

Shinn, M., Algozzine, B., Marston, D., & Ysseldyke, J. A theoretical analysis of the performance of learning disabled students on the Woodcock-Johnson Psycho-Educational Battery (Research Report No. 38). August, 1980.

Richey, L. S., Ysseldyke, J., Potter, M., Regan, R. R., & Greener, J. Teachers' attitudes and expectations for siblings of learning disabled children (Research Report No. 39). August, 1980.

Ysseldyke, J. E., Algozzine, B., & Thurlow, M. L. (Eds.). A naturalistic investigation of special education team meetings (Research Report No. 40). August, 1980.

Meyers, B., Meyers, J., & Deno, S. Formative evaluation and teacher decision making: A follow-up investigation (Research Report No. 41). September, 1980.

Fuchs, D., Garwick, D. R., Featherstone, N., & Fuchs, L. S. On the determinants and prediction of handicapped children's differential test performance with familiar and unfamiliar examiners (Research Report No. 42). September, 1980.

Algozzine, B., & Stoller, L. Effects of labels and competence on teachers' attributions for a student (Research Report No. 43). September, 1980.

Ysseldyke, J. E., & Thurlow, M. L. (Eds.). The special education assessment and decision-making process: Seven case studies (Research Report No. 44). September, 1980.

Ysseldyke, J. E., Algozzine, B., Potter, M., & Regan, R. A descriptive study of students enrolled in a program for the severely learning disabled (Research Report No. 45). September, 1980.

Marston, D. Analysis of subtest scatter on the tests of cognitive ability from the Woodcock-Johnson Psycho-Educational Battery (Research Report No. 46). October, 1980.

Algozzine, B., Ysseldyke, J. E., & Shinn, M. Identifying children with learning disabilities: When is a discrepancy severe? (Research Report No. 47). November, 1980.

Fuchs, L., Tindal, J., & Deno, S. Effects of varying item domain and sample duration on technical characteristics of daily measures in reading (Research Report No. 48). January, 1981.

Marston, D., Lowry, L., Deno, S., & Mirkin, P. An analysis of learning trends in simple measures of reading, spelling, and written expression: A longitudinal study (Research Report No. 49). January, 1981.

Marston, D., & Deno, S. The reliability of simple, direct measures of written expression (Research Report No. 50). January, 1981.

Epps, S., McGue, M., & Ysseldyke, J. E. Inter-judge agreement in classifying students as learning disabled (Research Report No. 51). February, 1981.

Epps, S., Ysseldyke, J. E., & McGue, M. Differentiating LD and non-LD students: "I know one when I see one" (Research Report No. 52). March, 1981.

Evans, P. R., & Peham, M. A. S. Testing and measurement in occupational therapy. A review of current practice with special emphasis on the Southern California Sensory Integration Tests (Monograph No. 15). April, 1981.

Fuchs, L., Wesson, C., Tindal, G., & Mirkin, P. Teacher efficiency in continuous evaluation of IEP goals (Research Report No. 53). June, 1981.

Fuchs, D., Featherstone, N., Garwick, D. R., & Fuchs, L. S. The importance of situational factors and task demands to handicapped children's test performance (Research Report No. 54). June, 1981.

Tindal, G., & Deno, S. L. Daily measurement of reading: Effects of varying the size of the item pool (Research Report No. 55). July, 1981.

Fuchs, L. S., & Deno, S. L. A comparison of teacher judgment, standardized tests, and curriculum-based approaches to reading placement (Research Report No. 56). August, 1981.

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

Christenson, S., Graden, J., Potter, M., & Ysseldyke, J. Current research on psychoeducational assessment and decision making: Implications for training and practice (Monograph No. 16). September, 1981.

Christenson, S., Ysseldyke, J., & Algozzine, B. Institutional constraints and external pressures influencing referral decisions (Research Report No. 58). October, 1981.

Fuchs, L., Fuchs, D., & Deno, S. Reliability and validity of curriculum-based informal reading inventories (Research Report No. 59). October, 1981.

Algozzine, B., Christenson, S., & Ysseldyke, J. Probabilities associated with the referral-to-placement process (Research Report No. 60). November, 1981.

Tindal, G., Fuchs, L. Christenson, S., Mirkin, P., & Deno, S. The relationship between student achievement and teacher assessment of short- or long-term goals (Research Report No. 61). November, 1981.

Mirkin, P., Fuchs, L., Tindal, G., Christenson, S., & Deno, S. The effect of IEP monitoring strategies on teacher behavior (Research Report No. 62). December, 1981.

Wesson, C., Mirkin, P., & Deno, S. Teachers' use of self instructional materials for learning procedures for developing and monitoring progress on IEP goals (Research Report No. 63). January, 1982.

Fuchs, L., Wesson, C., Tindal, G., Mirkin, P., & Deno, S. Instructional changes, student performance, and teacher preferences: The effects of specific measurement and evaluation procedures (Research Report No. 64). January, 1982.

Potter, M., & Mirkin, P. Instructional planning and implementation practices of elementary and secondary resource room teachers: Is there a difference? (Research Report No. 65). January, 1982.

Thurlow, M. L., & Ysseldyke, J. E. Teachers' beliefs about LD students (Research Report No. 66). January, 1982.

Graden, J., Thurlow, M. L., & Ysseldyke, J. E. Academic engaged time and its relationship to learning: A review of the literature (Monograph No. 17). January, 1982.

King, R., Wesson, C., & Deno, S. Direct and frequent measurement of student performance: Does it take too much time? (Research Report No. 67). February, 1982.

Greener, J. W., & Thurlow, M. L. Teacher opinions about professional education training programs (Research Report No. 68). March, 1982.

Algozzine, B., & Ysseldyke, J. Learning disabilities as a subset of school failure: The oversophistication of a concept (Research Report No. 69). March, 1982.

Fuchs, D., Zern, D. S., & Fuchs, L. S. A microanalysis of participant behavior in familiar and unfamiliar test conditions (Research Report No. 70). March, 1982.

Shinn, M. R., Ysseldyke, J., Deno, S., & Tindal, G. A comparison of psychometric and functional differences between students labeled learning disabled and low achieving (Research Report No. 71). March, 1982.

Thurlow, M. L. Graden, J., Greener, J. W., & Ysseldyke, J. E. Academic responding time for LD and non-LD students (Research Report No. 72). April, 1982.

Graden, J., Thurlow, M., & Ysseldyke, J. Instructional ecology and academic responding time for students at three levels of teacher-perceived behavioral competence (Research Report No. 73). April, 1982.

Algozzine, B., Ysseldyke, J., & Christenson, S. The influence of teachers' tolerances for specific kinds of behaviors on their ratings of a third grade student (Research Report No. 74). April, 1982.

Wesson, C., Deno, S., & Mirkin, P. Research on developing and monitoring progress on IEP goals: Current findings and implications for practice (Monograph No. 18). April, 1982.

Mirkin, P., Marston, D., & Deno, S. L. Direct and repeated measurement of academic skills: An alternative to traditional screening, referral, and identification of learning disabled students (Research Report No. 75). May, 1982.

Algozzine, B., Ysseldyke, J., Christenson, S., & Thurlow, M. Teachers' intervention choices for children exhibiting different behaviors in school (Research Report No. 76). June, 1982.

Tucker, J., Stevens, L. J., & Ysseldyke, J. E. Learning disabilities: The experts speak out (Research Report No. 77). June, 1982.

Thurlow, M. L., Ysseldyke, J. E., Graden, J., Greener, J. W., & Mecklenberg, C. Academic responding time for LD students receiving different levels of special education services (Research Report No. 78). June, 1982.

Graden, J. L., Thurlow, M. L., Ysseldyke, J. E., & Algozzine, B. Instructional ecology and academic responding time for students in different reading groups (Research Report No. 79). July, 1982.

Mirkin, P. K., & Potter, M. L. A survey of program planning and implementation practices of LD teachers (Research Report No. 80). July, 1982.

Fuchs, L. S., Fuchs, D., & Warren, L. M. Special education practice in evaluating student progress toward goals (Research Report No. 81). July, 1982.

Kuehnle, K., Deno, S. L., & Mirkin, P. K. Behavioral measurement of social adjustment: What behaviors? What setting? (Research Report No. 82). July, 1982.

Fuchs, D., Dailey, Ann Madsen, & Fuchs, L. S. Examiner familiarity and the relation between qualitative and quantitative indices of expressive language (Research Report No. 83). July, 1982.

Videen, J., Deno, S., & Marston, D. Correct word sequences: A valid indicator of proficiency in written expression (Research Report No. 84). July, 1982.

Potter, M. L. Application of a decision theory model to eligibility and classification decisions in special education (Research Report No. 85). July, 1982.

Greener, J. E., Thurlow, M. L., Graden, J. L., & Ysseldyke, J. E. The educational environment and students' responding times as a function of students' teacher-perceived academic competence (Research Report No. 86). August, 1982.

Deno, S., Marston, D., Mirkin, P., Lowry, L., Sindelar, P., & Jenkins, J. The use of standard tasks to measure achievement in reading, spelling, and written expression: A normative and developmental study (Research Report No. 87). August, 1982.

Skiba, R., Wesson, C., & Deno, S. L. The effects of training teachers in the use of formative evaluation in reading: An experimental-control comparison (Research Report No. 88). September, 1982.

Martson, D., Tindal, G., & Deno, S. L. Eligibility for learning disability services: A direct and repeated measurement approach (Research Report No. 89). September, 1982.

Thurlow, M. L., Ysseldyke, J. E., & Graden, J. L. LD students' active academic responding in regular and resource classrooms (Research Report No. 90). September, 1982.

Ysseldyke, J. E., Christenson, S., Pianta, R., Thurlow, M. L., & Algozzine, B. An analysis of current practice in referring students for psychoeducational evaluation: Implications for change (Research Report No. 91). October, 1982.

Ysseldyke, J. E., Algozzine, B., & Epps, S. A logical and empirical analysis of current practices in classifying students as handicapped (Research Report No. 92). October, 1982.

Tindal, G., Marston, D., Deno, S. L., & Germann, G. Curriculum differences in direct repeated measures of reading (Research Report No. 93). October, 1982.