DOCUMENT RESUME

ED 236 189 TM 830 698

AUTHOR King, Robert P.; And Others
TITLE The Effects of Training Teachers in the Use of Formative Evaluation in Reading: An Experimental-Control Comparison.
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-111
PUB DATE Feb 83
CONTRACT 300-80-0622
NOTE 33p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Criterion Referenced Tests; *Diagnostic Teaching; Elementary Education; *Elementary School Teachers; Evaluation Methods; *Formative Evaluation; Inservice Education; Measurement Techniques; *Program Effectiveness; Program Implementation; Reading Achievement; *Reading Instruction

ABSTRACT

A year long study involving 38 students in grades 1 to 6 was conducted to assess the degree of implementation of a frequent, curriculum-based measurement and evaluation system in classrooms in which the teachers had received training in the system, and to examine the effectiveness of the measurement and evaluation system in terms of enhancing the structure of the instructional lessons and students' reading achievement. The results indicated that although teachers were skillful in the measurement part of the system, they were unsuccessful in applying the evaluation components; students' instructional programs seldom were changed. In terms of the structure of the lessons, only one of the 12 structure variables (controlled practice) yielded significantly higher ratings for experimental than for control subjects. The remaining 11 variables favored experimental subjects, but were not statistically significant. No statistically significant differences in achievement were found between the two groups. All students improved over time. The results suggested that the implementation of a frequent curriculum-based measurement system is feasible and successful in improving the structure of instruction. Achievement effects may be manifest if the evaluation components are applied. (Author)

University of Minnesota

Research Report No. 111

THE EFFECTS OF TRAINING TEACHERS IN THE USE OF FORMATIVE EVALUATION IN READING: AN EXPERIMENTAL-CONTROL COMPARISON

Robert P. King, Stanley Deno, Phyllis Mirkin, and Caren Wesson

Institute for Research on Learning Disabilities

IRLD

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).

Research Report No. 111

THE EFFECTS OF TRAINING TEACHERS IN THE USE OF FORMATIVE EVALUATION IN READING: AN EXPERIMENTAL-CONTROL COMPARISON

Robert P. King, Stanley Deno, Phyllis Mirkin, and Caren Wesson

Institute for Research on Learning Disabilities

University of Minnesota

February, 1983

Abstract

A year long study involving 38 students was conducted to (a) assess the degree of implementation of a frequent, curriculum-based measurement and evaluation system in classrooms in which the teachers had received training in the system, and (b) examine the effectiveness of the measurement and evaluation system in terms of enhancing the structure of the instructional lessons and students' reading achievement. The results indicated that although teachers were skillful in the measurement part of the system, they were unsuccessful in applying the evaluation components; students' instructional programs seldom were changed. In terms of the structure of the lessons, only one of the 12 structure variables (controlled practice) yielded significantly higher ratings for experimental than for control subjects. The remaining 11 variables favored experimental subjects, but were not statistically significant. No statistically significant differences in achievement were found between the two groups. All students improved over time. The results suggested that the implementation of a frequent curriculum-based measurement system is feasible and successful in improving the structure of instruction. Achievement effects may be manifest if the evaluation components are applied.

The Effects of Training Teachers in the Use of Formative Evaluation in Reading: An Experimental-Control Comparison

In recent years, with the advent of Public Law 94-142 and increased public pressure for accountability in education, greater demands have been placed on educators, especially special educators, to be accountable for the quality of instructional decisions and the ways in which they are made. Recent evidence (White & Haring, 1980) suggests that formative evaluation systems may provide viable alternatives to the traditional pre and post testing approach to evaluation of academic programs. Such systems provide continuous feedback to both the teacher and student, allowing educators to more closely monitor academic progress.

During the past five years, the Institute for Research on Learning Disabilities at the University of Minnesota, under federal contract, has conducted a number of studies that focused on developing and monitoring progress on IEP goals, as is intended in PL 94-142. The goal of this research has been to determine empirically the effects of using formative evaluation techniques on student achievement in reading, spelling, and written expression.

Earlier research in this area determined what measures of student performance would be ideal for use in a formative evaluation system. The search for these measures began by generating a list of desired characteristics, such as ease of administration, time efficiency, and sensitivity to growth over time (Jenkins, Deno, & Mirkin, 1979). The measures that were not reliable or valid, or those that were deemed less suitable with respect to any of the other desired characteristics, were eliminated from consideration.

Five reading behaviors were generated from a review of the literature and placed in the original pool for consideration. A series of criterion validity studies (Deno, Mirkin, Chiang, & Lowry, 1980), showed that reading aloud from a basal reader, reading aloud from lists of isolated words, and guessing the words deleted from a reading passage (i.e., cloze comprehension) all related closely to performance on standardized tests and discriminated between program and grade placement. Such formative measures of reading have also shown high test-retest (r = .90) and alternate forms (rs = .89 - .92) reliability (Shinn, 1981).

Related studies focused on determining the optimal duration of reading measurement and the type of data to record. Results from testing over one, two, and three-minute durations indicated that reading proficiency can be indexed validly within one minute and that correct performance is a more valid measure of reading proficiency than error performance (Deno et al., 1980).

Previous studies also assessed the sensitivity of two reading measures, reading isolated word lists and reading aloud from a basal reader. Both reading measures were found to be sensitive to changes within each grade level from fall to spring and across grade levels (Marston et al., 1981). However, reading aloud from a basal reader was chosen as the optimal generic measure in reading because it produced a broader range of scores than isolated words, related somewhat more closely to comprehension, and required little teacher preparation.

Given that one-minute timed samples of reading from the

curriculum have been shown to be reliable and valid measures of reading growth, there remained the need to test the practicality of such measures and the effects that teacher use of such measures might have on student achievement over time. Specific questions related to these issues were posed in the current study.

First, can teachers learn to use the measurement system and will they find it practical and time-efficient? Once the measurement system is implemented, will teachers use the information it provides to more closely monitor and change the educational program of the student? One of the major advantages of such a system is that it allows for continuous evaluation of the instructional program. Thus, it is critical that the information provided by the system be used.

Only if these questions are answered affirmatively is it possible to examine the questions concerning the effectiveness of the system. Two questions were investigated concerning the efficacy of the measures. First, will the use of such measures have an effect on the structure of the learning environment provided to the student? Because a formative evaluation system provides continuous information about the need for program changes, one might expect the use of such a system to result in a more highly structured learning environment. Second, given that teachers can learn to use such a formative evaluation system for both measurement and evaluation, will the use of such procedures have a direct effect on student achievement? One would expect that frequent modifications in the instructional plan made possible by continuous feedback would lead to an educational program more sensitive to individual needs and thus more conducive to

growth in reading.

Method

Subjects

A total of 38 elementary students in grades 1-6 participated in the study. (See Table 1 for complete breakdown of subjects by grade and sex.) Thirty-two of the students were male (84.2%) and six were female (15.8%). Students were assigned to either the experimental (treatment) or control (no treatment) conditions for comparison purposes. Data were obtained on 19 students in each condition.

Insert Table 1 about here

Procedures

Designated trainers from a large midwest suburban school district participated in a full-day workshop before the beginning of the school year. Principally, training focused on the use of the measurement and evaluation procedures as prescribed in the IRLD manual entitled Procedures to Develop and Monitor Progress on IEP Goals (Mirkin, Deno, Fuchs, Wesson, Tindal, Marston, & Kuehnle, 1981). Subsequent to this workshop, the trainers trained the teacher participants in the use of the measurement and evaluation procedures.

Daily measurement consisted of one-minute timed samples of reading from the student's curriculum. Both words correct and incorrect were scored and graphed on equal interval charts. Based on the results of previous research (Fuchs & Deno, 1981), the placement level for testing, which also became the baseline, was set at a

criteria of 20-29 words per minute for grades 1 and 2, and 30-39 words per minute for grades 3 through 6.

Teachers were instructed to write IEP long-range goals using both the entry level criteria and a desired year-end mastery criteria, usually 70 words correct per minute with no more than 7 errors.

Short-term objectives were based on the long-range goals (LRG). In order to compute the short-term objective, teachers first subtracted the baseline level of performance from the criterion level listed in the LRG. Dividing this difference by the number of weeks necessary until the annual review, they arrived at the number of words per week gain necessary to meet the long-range goal criteria.

In order to monitor student growth, the baseline reading level and the long-range goal were connected by an aimline that showed the students' desired progress. Every seven data points, the teachers were to monitor student growth by means of the split-middle or quarter-intersect method (White & Liberty, 1976). If the student was progressing at a rate equivalent to or greater than that indicated by the aimline, the instructional program was continued; if the projected rate of growth was less than that indicated by the aimline, teachers were directed to make a substantial change in the student's program.

Measures

Four measures were used in collecting data: one each for implementation and structure, and two for achievement. The structure of the learning environment was assessed by means of the Structure of Instruction Rating Scale for both experimental and control subjects (Deno, King, Skiba, Sevcik, & Wesson, 1983). Degree of implementation

of the continuous evaluation measures--the treatment for the experimental subjects-was assessed using the Accuracy of Implementation Rating Scale. Achievement measures for both experimental and control groups consisted of timed samples from three third grade passages, and subtests of the Stanford Diagnostic Reading Test (SDRT). The three timed samples were collected three times during the year. The Stanford Diagnostic Reading Test was administered only once, in May, to both experimental and control subjects. Descriptions of the measures follow.

Structure of instruction rating scale. The Structure of Instruction Rating Scale (SIRS) was designed to measure the degree of structure of the instructional lesson that a student received, in this case in reading. The variables chosen for inclusion on the SIRS were gathered from current literature on instruction and student academic achievement (cf. Stevens & Rosenshine, 1981).

The SIRS consists of 12 five-point rating scales in which a rating of 1 is low for the variable and 5 is high. The reliability of the SIRS was assessed by means of Coefficient Alpha, a measure of internal consistency. For a sample of 70 students observed in November, the average inter-item correlation was .37, resulting in an alpha of .86. Thus, the SIRS seems to have a high degree of reliability as indexed by measures of homogeneity.

Factor analysis of the 12 variables on the SIRS revealed that 9 of the 12 represented one factor. Three variables--Independent Practice, Positive Consequences, and Silent Practice on Outcome Behavior--were not measuring the same factor. Thus, the nine variables were utilized

in the data analyses as one factor and the other three variables were analyzed separately.

Accuracy of implementation rating scale. The Accuracy of Implementation Rating Scale (AIRS) is an instrument that was developed in conjunction with the manual Procedures to Develop and Monitor Progress on IEP Goals (Mirkin et al., 1981). The AIRS is designed to provide a format by which to monitor the implementation of the procedures described in the manual. The AIRS consists of 12 items that are rated on a 1 to 5 scale, 1 being the lowest implementation score and 5 being complete and accurate implementation.

Parts of the scale require direct observation whereas other items on the checklist can be monitored by inspection of student reading graphs and by reading IEP forms. Items 1 and 2 of the AIRS, which require direct observation, deal with the accuracy of administration of the measurement system and selection of the stimulus materials. For items 3-12 of the AIRS, research assistants inspected various written documents and made the ratings. Specifically, the IRLD rater examined the following documents for each student: (a) the IEP, which should specify the long-range goal and short-term objective in reading; (b) the reading graph; (c) the instructional plan for reading; and (d) the record of changes made in the instructional plan in reading. Factors included in items 3-12 pertain to the establishment of the appropriate measurement level, an adequate baseline, an accurate long-range goal and short-term objective, a detailed graph, a complete instructional program, and a correct aimline. These items also included the timing of instructional

changes and the types of changes made. Frequent checks among the four research assistants rating the accuracy of implementation assured high inter-rater agreement. Reliability of the AIRS was assessed by means of the Cronbach's Alpha internal consistency measure. The average inter-item correlation was .12, resulting in an alpha of .62.

Results

Implementation (AIRS)

The mean raw score ratings for each variable on the AIRS for each round of data collection for the experimental students are reported in Table 2. As mentioned previously, variables were rated on a 1 (low) to 5 (high) rating scale. Ratings were assessed by IRLD staff for all of the AIRS variables except variables 1 and 2, which were scored by district observers/trainers.

Insert Table 2 about here

The data strongly indicate that during all three rounds of data collection, teachers consistently were able to employ the initial measurement procedures, e.g., administering the measurement task, selecting the stimulus materials, obtaining a baseline measure of performance, labeling the graph appropriately, determining short term and long range objectives, as well as determining the aimline based on a formula outlined in the manual Procedures for Developing and Monitoring Progress on IEP Goals (Mirkin et al., 1981). Areas where teacher implementation scores could be higher involved evaluation and utilization of the data on an on-going basis. For example, Timing of

Instructional Changes, Substantial Changes, and Clear Changes were rated considerably lower than the other AIRS variables previously discussed. For this sample, only 10 changes in the instructional plan were recorded over a five-month period. In general, teachers seldom changed the instructional plan once it was established.

Structure of Instruction (SIRS)

The mean ratings for each variable and *t* values for the experimental and control group comparisons are reported in Tables 3, 4, and 5. The data indicated that nine variables consistently stayed together across the three time conditions; these emerged as a separate factor. The moderate to high ratings on these variables at all three points in time suggest that these aspects of classroom structure are fairly stable and present in the classrooms. Interestingly, statistically significant differences between experimental and control groups were recorded for the variable Controlled Practice across all three times. At Time One, the control students received significantly higher ratings on this variable than the experimental students. However, for both Time Two and Time Three the higher mean ratings for the experimental students were statistically significant. For both groups of students, the variables Independent Practice, Positive Consequences, and Silent Practice on Outcome Behavior were rated considerably lower than the other variables constituting classroom structure.

Insert Tables 3-5 about here

Controlled Practice was the only variable for which differences in sample means between the control and experimental students was statistically significant. At Time Three, the sample mean ratings for 9 of the 12 SIRS variables were higher for the experimental students, but none of the differences was statistically reliable.

Achievement

Data on the number of correct words read per minute on each of three reading passages are reported in Table 6. Data included in the table were standardized to z scores using data from students in three additional research sites. Using a large normative sample to standardize scores increases the validity of the data and adjusts for the relatively low frequency of cases reported at grades 1 and 2 and 6. For ease of presentation, these z scores have been transformed into t scores using a standard deviation of 10 and a mean of 100. Results from an analysis of variance (see Table 7) indicated that all students, on the average, showed growth over time. However, the gains for the experimental students were not significantly different from those for the control students.

Insert Tables 6 and 7 about here

Raw score data for the various subtests on the Stanford Diagnostic Reading Test (SDRT) for both experimental and control students are reported in Table 8. Analyses by t-test comparisons revealed no statistically significant differences between the two groups. However, the sample means were slightly higher for the

control students on all six SDRT subtests.

Insert Table 8 about here

Discussion

The present investigation focuses on a number of important questions relating to the practicality and effects of direct and frequent monitoring of progress on IEP goals. Principally, can teachers learn to use such a measurement system? Additionally, will teachers use the information provided by such a system to make frequent changes in the educational plan and monitor the effectiveness of those changes? Moreover, will such a system have an effect on the structure of reading instruction the student receives and will this be related to reading achievement?

Data from the present investigation revealed that teachers can learn to effectively administer timed reading samples and accurately chart the data to provide a continuous record of student growth in reading. Ratings of trainers on the Accuracy of Implementation Rating Scale support this finding. However, data also indicated that the teachers only partially used the evaluation component of the data-based system. That is, the teachers' use of procedures to evaluate student data in order to make on-going changes in the instructional plan was low; teachers could have made use of these procedures to a considerably greater extent.

Also, data generated to assess the structure of the instructional lesson revealed that experimental students engaged in significantly

more controlled practice of their lessons than the control students. Moreover, by Time Three, structure variables more often were rated higher for the experimental students. This finding suggests that teachers who utilized the data based system provided greater structure for the reading lessons. However, some structure variables were consistently rated lower for both the control and experimental groups. For example, Positive Consequences rarely was an aspect of the classroom setting. Given that many of the students in this sample had difficulties in reading, it is surprising that some form of contingency management or token economy was not used more often as a motivator for improving the reading performance of these students.

Usefulness of the Procedures

At the end of the year, teachers who participated in the study completed questionnaires regarding their reactions to the data based program modification procedures. These data currently are being analyzed as part of a larger study. However, preliminary findings are quite favorable. Moreover, data gathered informally during a presentation of findings at the end of the year suggest that both trainers and teachers, for the most part, believed that the system provides an indication of reading progress and growth.

Although the present study did not support the contention that teachers can use the evaluation system effectively to increase reading achievement, the results do demonstrate the feasibility of using such a system to monitor progress on IEP goals routinely--a necessary component of special education programs (PL 94-142). While teachers in the present sample were not successful in using the evaluation

components of the system, preliminary findings from a similar, though larger scale, experiment in the New York City public schools support the efficacy of using such an approach to monitor and evaluate reading progress (Fuchs, Deno, & Mirkin, 1982).

References

Deno, S. L., King, R., Skiba, R., Sevcik, B., & Wesson, C. The structure of instruction rating scale (SIRS): Development and technical characteristics (Research Report No. 107). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1983.

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). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1980. (ERIC Document Reproduction Service No. ED 197 507)

Fuchs, L. S., & Deno, S. L. A comparison of reading placements based on teacher judgment, standardized testing, and curriculum-based assessment (Research Report No. 56). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1981. (ERIC Document Reproduction Service No. ED 211 603)

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. Effects of frequent curriculum-based measurement on student achievement and knowledge of performance: An experimental study (Research Report No. 96). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1982.

Jenkins, J. R., Deno, S. L., & Mirkin, P. K. Measuring pupil progress toward the least restrictive environment (Monograph No. 10). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1979. (ERIC Document Reproduction Service No. ED 185 767)

Marston, D., Lowry, L., Deno, S. L., & Mirkin, P. K. An analysis of learning trends in simple measures of reading, spelling, and written expression: A longitudinal study (Research Report No. 49). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1981. (ERIC Document Reproduction Service No. ED 211 602)

Mirkin, P., Deno, S., Fuchs, L., Wesson, C., Tindal, G., Marston, D., & Kuehnle, K. Procedures to develop and monitor progress on IEP goals. Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1981.

Shinn, M. A comparison of psychometric and functional differences between students labelled learning disabled and low achieving. Doctoral dissertation, University of Minnesota, 1981.

Stevens, R., & Rosenshine, B. Advances in research on teaching. Exceptional Education Quarterly, 1981, 2(1), 1-10.

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

White, O. R., & Liberty, K. A. Behavioral assessment and precise educational measurement. In N. G. Haring & R. L. Schiefelbusch (Eds.), Teaching special children. New York: McGraw-Hill, 1976.

Table 1

Grade, Sex, and Age of Students

	Number of Students	Percentage
Grade		
1	1	2.6%
2	2	5.3%
3	6	15.8%
4	11	28.9%
5	10	26.3%
6	5	13.2%
Unknown	3	7.9%
Total	38	100.0%
Sex		
Male	32	84.2%
Female	6	15.8%
Total	38	100.0%
Age (yrs)		
7	2	5.3%
8	3	7.9%
9	4	10.5%
10	11	28.9%
11	9	23.7%
12	5	13.2%
13	1	2.6%
Unknown	3	7.9%
Total	38	100.0%

Table 2

Mean Scores on the Accuracy of Implementation Rating Scale (AIRS)[a]

	Time 1	Time 2	Time 3
1. Administering the Measurement Task	4.00	4.76	4.45
2. Selecting the Stimulus Material	3.86	4.76	4.45
3. Sampling for Instructional Level	4.00	3.35	3.33
4. Baseline	4.29	3.68	3.86
5. Graph Set-up	4.00	3.82	3.93
6. Aimline	4.47	4.94	4.60
7. Timing of Instructional Changes	0.00	3.23	3.00
8. Long-Range Goal	4.70	4.82	4.73
9. Short-Term Objective	3.70	3.94	3.33
10. Instructional Plan	3.80	3.11	3.53
11. Substantial Changes	0.00	2.25	2.00
12. Clear Change	0.00	3.50	1.25

[a]Data are for experimental subjects only (N=19). Rating scale: 1=low, 5=high.

Table 3

Mean Scores on the Structure of Instruction Rating Scale (SIRS) and t Test Results for Time 1

		Mean	Separate Error Variance t	df	2-tail Prob
Instructional Grouping	E	4.38			
	C	4.06	1.15	26.45	.26
Teacher-directed Learning	E	4.11			
	C	4.13	-.07	30.42	.95
Active Academic Responding	E	4.27			
	C	4.20	.28	30.53	.78
Demonstration/Prompting	E	3.83			
	C	4.00	-.51	27.81	.62
Controlled Practice	E	3.44			
	C	4.21	-2.80	29.90	.01*
Frequency of Correct Answers	E	4.16			
	C	4.33	-.77	29.96	.45
Independent Practice	E	2.33			
	C	3.28	-1.28	9.11	.23
Corrections	E	4.11			
	C	4.13	-.07	27.25	.94
Positive Consequences	E	2.66			
	C	2.33	.73	30.23	.47
Pacing	E	4.27			
	C	4.13	.43	30.79	.67

Table 4

Mean Scores on the Structure of Instruction Rating Scale (SIRS) and t Test Results for Time 2

		Mean	Separate Error Variance t	df	2-tail Prob
Instructional Grouping	E	3.94			
	C	3.55	.95	34.60	.35
Teacher-directed Learning	E	4.00			
	C	3.77	.80	34.43	.43
Active Academic Responding	E	4.31			
	C	4.00	1.13	34.00	.27
Demonstration/Prompting	E	3.94			
	C	3.76	.62	32.00	.54
Controlled Practice	E	3.94			
	C	3.25	2.28	29.16	.03*
Frequency of Correct Answers	E	4.15			
	C	4.27	-.51	34.29	.62
Independent Practice	E	1.87			
	C	2.50	-.76	9.62	.47
Corrections	E	4.21			
	C	3.66	1.80	31.84	.08
Positive Consequences	E	2.78			
	C	2.22	1.23	34.56	.23
Pacing	E	3.94			
	C	3.66	.76	34.53	.45
Oral Practice on Outcome Behavior	E	3.68			
	C	3.38	.70	34.75	.49
Silent Practice on Outcome Behavior	E	2.36			
	C	2.77	-.87	35.00	.39

Table 5

Mean Scores on the Structure of Instruction Rating Scale (SIRS) and t Test Results for Time 3

		Mean	Separate Error Variance t	df	2-tail Prob
Instructional Grouping	E	4.31			
	C	3.76	1.65	30.87	.11
Teacher-directed Learning	E	4.05			
	C	3.88	.55	33.42	.58
Active Academic Responding	E	4.31			
	C	4.11	.69	33.80	.50
Demonstration/Prompting	E	4.26			
	C	4.05	.75	33.34	.46
Controlled Practice	E	3.89			
	C	3.11	2.17	28.74	.04*
Frequency of Correct Answers	E	4.00			
	C	4.05	-.21	33.91	.84
Independent Practice	E	2.22			
	C	2.66	-1.11	14.62	.28
Corrections	E	4.10			
	C	3.88	.65	31.49	.52
Positive Consequences	E	2.42			
	C	1.94	1.11	33.96	.27
Pacing	E	4.00			
	C	3.88	.29	33.92	.77
Oral Practice on Outcome Behavior	E	3.68			
	C	3.58	.22	33.95	.83
Silent Practice on Outcome Behavior	E	2.31			
	C	2.47	-.33	33.30	.74

Table 6

Reading Passage Data: T Score Transformation and Analysis of Variance

Reading Passage	Group	Time 1	T Scores	Time 2	T Scores	Time 3	T Scores
3	E	-.6969	93.1	-.3398	96.6	-.4585	95.4
	C	-.1158	98.9	-.0555	99.9	-.0561	99.4
4	E	-.7163	92.9	-.2261	97.7	-.4832	95.2
	C	-.3076	96.9	-.0645	99.4	-.1324	98.7
5	E	-.6704	93.3	-.3074	96.9	-.1547	98.5
	C	-.3054	97.0	-.2576	97.4	0841	99.2

Table 7

Analysis of Variance Results for

Reading Passage Data

Passage	df	F	prob
3	2,21		
Time		5.16	.02*
Time X Cond		1.21	.32
4	2,19		
Time		7.66	.00*
Time X Cond		0.50	.61
5	2,19		
Time		1.65	.22
Time X Cond		0.21	.81

Table 8

Mean Raw Scores on Subtests of the Stanford Diagnostic Reading (SDRT) and t Test Results

SDRT Subtest		Mean	Separate Error Variance t	df	2-tail Prob
Word Division	E	21.40			
	C	23.30	-1.06	33.14	.30
Word Blending	E	19.42			
	C	21.33	-.87	33.69	.39
Structural Analysis	E	40.84			
	C	44.66	-1.03	33.72	.31
Literal Comprehension	E	21.55			
	C	23.18	-.59	26.73	.56
Inferential Comprehension	E	18.27			
	C	20.62	-.85	26.74	.40
Comprehension Total	E	39.83			
	C	43.81	-.74	26.67	.47

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 $4.00 each, 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.

The publications listed here are only those that have been prepared since 1982. For a complete, annotated list of all IRLD publications, write to the Editor.

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.

Marston, 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.

Fuchs, L.S., Deno, S. L., & Marston, D. Use of aggregation to improve the reliability of simple direct measures of academic performance (Research Report No. 94). October, 1982.

Ysseldyke, J. E., Thurlow, M. L., Mecklenburg, C., & Graden, J. Observed changes in instruction and student responding as a function of referral and special education placement (Research Report No. 95). October, 1982.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. Effects of frequent curriculum-based measurement and evaluation on student achievement and knowledge of performance: An experimental study (Research Report No. 96). November, 1982.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. Direct and frequent measurement and evaluation: Effects on instruction and estimates of student progress (Research Report No. 97). November, 1982.

Tindal, G., Wesson, C., Germann, G., Deno, S. L., & Mirkin, P. K. The Pine County model for special education delivery: A data-based system (Monograph No. 19). November, 1982.

Epps, S., Ysseldyke, J. E., & Algozzine, B. An analysis of the conceptual framework underlying definitions of learning disabilities (Research Report No. 98). November, 1982.

Epps, S., Ysseldyke, J. E., & Algozzine, B. Public-policy implications of different definitions of learning disabilities (Research Report No. 99). November, 1982.

Ysseldyke, J. E., Thurlow, M. L., Graden, J. L., Wesson, C., Deno, S. L., & Algozzine, B. Generalizations from five years of research on assessment and decision making (Research Report No. 100). November, 1982.

Marston, D., & Deno, S. L. Measuring academic progress of students with learning difficulties: A comparison of the semi-logarithmic chart and equal interval graph paper (Research Report No. 101). November, 1982.

Beattie, S., Grise, P., & Algozzine, B. Effects of test modifications on minimum competency test performance of third grade learning disabled students (Research Report No. 102). December, 1982

Algozzine, B., Ysseldyke, J. E., & Christenson, S. An analysis of the incidence of special class placement: The masses are burgeoning (Research Report No. 103). December, 1982.

Marston, D., Tindal, G., & Deno, S. L. Predictive efficiency of direct, repeated measurement: An analysis of cost and accuracy in classification (Research Report No. 104). December, 1982.

Wesson, C., Deno, S., Mirkin, P., Sevcik, B., Skiba, R., King, R., Tindal, G., & Maruyama, G. Teaching structure and student achievement effects of curriculum-based measurement: A causal (structural) analysis (Research Report No. 105). December, 1982.

Mirkin, P. K., Fuchs, L. S., & Deno, S. L. (Eds.). Considerations for designing a continuous evaluation system: An integrative review (Monograph No. 20). December, 1982.

Marston, D., & Deno, S. L. Implementation of direct and repeated measurement in the school setting (Research Report No. 106). December, 1982.

Deno, S. L., King, R., Skiba, R., Sevcik, B., & Wesson, C. The structure of instruction rating scale (SIRS): Development and technical characteristics (Research Report No. 107). January, 1983.

Thurlow, M. L., Ysseldyke, J. E., & Casey, A. Criteria for identifying LD students: Definitional problems exemplified (Research Report No. 108). January, 1983.

Tindal, G., Marston, D., & Deno, S. L. The reliability of direct and repeated measurement (Research Report No. 108). February, 1983.

Fuchs, D., Fuchs, L. S., Dailey, A. M., & Power, M. H. Effects of pre-test contact with experienced and inexperienced examiners on handicapped children's performance (Research Report No. 110). February, 1983

King, R. P., Deno, S., Mirkin, P., & Wesson, C. The effects of training teachers in the use of formative evaluation in reading: An experimental-control comparison (Research Report No. 111). February, 1983.