DCCUMENT RESUME

ED 049 902 RE 003 574

AUTHCR	Quilling, Mary F.
TITLE	The Reading Achievement of Primary Age Pupils Using the Wisconsin Design for Reading Skill Develcpment: A Ccmparative Study.
PUB DATE	Feb 71
CONTRACT	OEC-5-10-154
NOTE	15p.; Paper presented at the meeting of the American Educational Research Asscciation, New York, N.Y., Feb. 4-7, 1971
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Beginning Reading, Ccnference Reports, Elementary School Students, Grade 1, Grade 2, Grade 3, Primary Grades, *Program Evaluation, Reading Achievement, *Reading Instruction, Reading Programs, Reading Research, *Research and Development Centers, *Structural Analysis, Word Recognition

ABSTRACT

Prepared by the Wisconsin Research and Develcpment Center for Ccgnitive Grcwth, the Wisconsin Design for Reading Skill Development (Design) contains several components. The field study evaluation of the Word Attack element in terms of attainment of objectives is reported in this ccnference paper. All children in grades 1 to 3 of two Wisccnsin schools participated in the program evaluation during the 1969-70 school year. They were tested at the beginning and at the end of the prcgram using design-developed criterion-referenced tests and selected subtests of the Doren Diagnostic Reading Test. Both tests registered greater gains for students who had Design instruction over those who had not. In School A, where the Stanford Achievement Test is used, no gains were noted for the Desigr group and possible reasons for this are discussed. In School B, where the Gates-MacGinitie Reading Test is used, greater gains were evident for the Design group. Tables of results are included. (MS)

THE READING ACHIEVEMENT OF PRIMARY AGE PUPILS USING THE WISCONSIN DESIGN FOR READING SKILL DEVELOPMENT: A COMPARATIVE STUDY

Mary R. Quilling

Wisconsin Research and Development
Center for Cognitive Learning

A paper presented at the annual meeting of the
American Educational Research Association.
February 4-7, 1971 New York City

The material reported herein was prepared with the support of the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

THE READING ACHIEVEMENT OF PRIMARY AGE PUPILS USING THE WISCONSIN DESIGN FOR READING SKILL DEVELOPMENT: A COMPARATIVE STUDY

MARY R. QUILLING
WISCONSIN RESEARCH AND DEVELOPMENT
CENTER FOR COGNITIVE LEARNING

The Wisconsin Design for Reading Skill Development (Design) is a product of the Wisconsin Research and Development Center for Cognitive Learning. Like other Center products it is evaluated in terms of specifications and objectives established at the outset of the developmental effort. Each component is subjected to expert review and subsequently to empirical validation through a series of field tests. The principal purpose of each of the field tests is to determine whether or not the objectives of the product are attained when implementation is carried out according to plan.

During the first field study or pilot, monitoring and process evaluation lead to modifications in specific aspects of the materials and procedures. From information gathered during the pilot the developer is able to decide how to proceed in revising the product prototype, and whether to move on or to iterate in the development sequence. Some of the data collected at this point in the evaluative process, however, may be regarded as summative in nature. If the pilot has been conducted under fairly typical school conditions and if only a few minor modifications are required in the materials, the evaluator may wish to suggest that the summative evaluation of the product was in fact beginning during the formative evaluation period.

Formative evaluation of the Word Attack element of the Design was carried out at the primary levels (grades 1 to 3) in two schools during 1969-1970. Summative data were also collected at this time.

The pilot population. All pupils in both schools who were in their second, third or fourth year of school participated in the reading program during the 1969-1970 school year. Both schools are in predominantly white neighborhoods in small Wisconsin cities. Mean IQ, as measured in the third year of school by the Kuhlmann-Anderson Intelligence test, is 111 for School A and 100 for School B. School B has a high proportion of pupils from broken homes and whose mothers receive welfare payments.

Product objectives and instrumentation. The terminal outcome anticipated for pupils participating in the Word Attack program is as follows:

> The student upon attainment of all Level D Skills will be able to attack independently, phonically and/or structurally regular words and will recognize on sight all the words on the Dolch list. Children of average or above average ability will attain this objective at least by the end of the fifth year (fourth grade) in school, while others will attain this objective by the end of the seventh year.

It is presumed that this outcome will be realized if participants attain at a steady rate the 45 specified objectives which serve as a framework for the program. These objectives are behaviorally stated and are arranged into four levels (A through D). While the Word Attack program will eventually be evaluated in relation to its terminal objective during the first year of implementation evaluation is carried out in relation to the specific objectives.

Criterion-referenced pencil and paper tests had been constructed for 36 of the 45 objectives at the time of the study; attainment of the remaining objectives was assessed either by individually administered tests involving the pronunciation of words or by teacher observation. The child breaks into the program by taking the set of tests at the level his teacher believes is most appropriate for him. If he wholly fails or succeeds at a given level he is administered the battery at the next lower or higher level. Instructional programming associated with the Design then calls for three-week skill groupings of children with common deficiencies.

During the three-week period the appropriate criterion-referenced assessment procedure is administered when an individual child may, in the judgment of his teacher, be ready to demonstrate his mastery of the objective. The immediate effects of the program, then, are readily observable in terms of pupil attainment of objectives.

There are several difficulties inherent in relying upon the pupils' skill profile at a given moment as the source of information for assessing attainment of program objectives: 1) the conditions under which the test was administered may not always meet the evaluator's standards; 2) the scoring and record keeping are subject to human error; 3) skill mastery over the long term, not just immediately following instruction, is of interest; and 4) the practice effects of administering the same test more than once might account for any positive results.

For these reasons tests referenced to objectives in the Design were readministered as part of the evaluation procedure. These tests were given in both schools at the beginning of instruction in September 1969 and again, for evaluation purposes, in School A during September 1970. To confirm the results of the program-related testing program, a program-independent test was also given. Selected subtests of the Doren Diagnostic Reading Test were selected for this purpose because the content of these subtests was similar to a number of the skills in the lower levels of the program. This test was administered in both schools in May, 1969 and May, 1970 to children completing their third year (Grade 2).

While the primary purpose of the evaluation was to determine that the specific program objectives were attained, a secondary objective was to explore the effect of the program on general reading achievement. The standardized testing program used in each school was implemented as required

by the district, and the data made available to the Center. Different standardized tests were administered in the two schools, which were located in different school districts.

Table 1 summarizes the schedule of data collection in each school. In all instances tests were administered to children of a particular age/grade group both in 1969 and in 1970, enabling comparisons to be made. The 1969 data, gathered in May, September or December, were pre-implementation baseline data, whereas data collected in 1970 were gathered five months to one year and two months after the program was initially implemented from children who had experienced the program. The data collection schedule is cumbersome for our purposes; nonetheless, it is justifiable because of its utilization of data necessarily collected for instructional and other evaluative purposes.

The results. Results from two administrations of the criterion-referenced tests one year apart are of primary importance for evaluating the attainment of program objectives. The data may be analyzed in two ways. First, the prior-to-implementation performance of children of a particular age/grade group may be compared one year later with the post-implementation performance of a different group of children of the same age/grade characteristics. For instance, children beginning their third year of school who have not used the program are compared one year later with beginning third year students who participated in the program during their second year of school. Another use of the data involves following the same group of children from one year to the next to determine the gain in performance. In each instance, the information of interest is computed from the pupil x skill matrix for each group in which dichotomous mastery/non-mastery data are entered.

In Table 2 data by which different groups of children can be compared are presented. For Levels A through C, pencil and paper tests were used

during the study for 30 out of 38 skills. For 23 of these skills, the percent of children demonstrating mastery was greater for the groups which had participated in the program than it was for the groups which had not. For one skill there was no difference in the performance of the two groups; for two of the six skills in which a negative effect was observed for program participants, different tests were used for the two groups, making the comparison inconclusive.

The distribution of gains in number of skills mastered for participants during their second, third and fourth years of school is presented in Table 3. Median gains of 8, 19 and 11 skills respectively were observed for the three age/grade groups. If a child were to attain about five skills per semester beginning with the second semester of Kindergarten, he would complete the program in the time projected by the developer. Also, the three-week skill groupings called for in the instructional programming model suggest that the child typically will have an opportunity to attain about 12 skills annually, if a single skill is acquired in each of the ad hoc groupings. The uneven distribution of gains across age/grade groups is apparently explained in several ways. First, certain of the skills beginning readers must acquire, such as letter-sound correspondences, require more than a single three-week session. Secondly, the child in his third year frequently has an opportunity to acquire more than one skill every three weeks if the skills are clustered for instruction as recommended in the manual. Finally, ceiling effects are noted in the fourth year as children who had most of the skills in their repertoire at the outset of program implementation have the opportunity to acquire only a few additional skills.

The results of two administrations of selected subtests from the Doren Diagnostic Reading Test in both schools are found in Table 4. In

School B the mean difference of 6.3 score points was highly significant ($p < .01$) and in School A the observed difference of 3.1 points was marginally significant ($p < .20$) in favor of the groups participating in the program. The ceiling effects which might have been anticipated for a diagnostic test were observed in both schools, particularly in the second year, and were especially acute in the school which realized the smaller gain. For most of the subtests a positive increment in performance was associated with program implementation.

Analyses of standardized achievement test data gathered in two successive years in the two schools, are inconclusive. In School A performance on the Word Study Skills subtest of the Stanford Achievement tests is of special interest. As indicated in Table 5, no difference was observed in median performance of the comparison groups at either age/grade level in this subtest. Shifts in the distribution of scores from year to year were minor; those observed, however, were slightly negative for the third year groups, and slightly positive in the fourth year group. Performance of the groups participating in the program was lower on the remaining subtests than was performance of the baseline groups. This outcome is in part attributable to the focus on word attack in the initial year of implementation. Introduction of the comprehension element of the Design may be expected to improve pupil performance on at least the paragraph meaning subtest.

In School B more uniformly positive results were observed when the performance of children on a standardized test administered in two successive years was compared. In School B the Gates-MacGinitie Reading Test was administered to all pupils completing their second, third or fourth year of school in 1969 or in 1970. Like the mastery information presented

earlier, two comparisons may be made: that of different groups of children of the same age/grade designation in successive years, and that indicating growth of a particular group of children from one year to the next. From the data in Table 6, one may conclude that there was a noticeable positive shift in the distribution of performance in the second administration of the test at each age/grade level on both tests; five of the six medians were higher for the 1970 test administration, first quartile scores were as high or higher in 1970 as in 1969, and all third quartile scores increased, some dramatically. The greater spread in the distribution of scores is an outcome one might anticipate with proper implementation of an individualized program.

When the gains made in the course of a year of program implementation are extracted from the data, one observes improvements of a year or more at the median and third quartile points. As might be expected the amount of gain is related to the point in the distribution one is considering. The first quartile gains for children during the fourth year of school are noteworthy for their magnitude, as are the year or better gains at the median point in a school where typical performance is often below grade level.

Summary and conclusions. The Word Attack element of the Wisconsin Design for Reading Skill Development was evaluated in terms of pupil attainment of objectives. Pupils attained a reasonable number of objectives in a year's time; also, for 23 of 30 skills the percent of pupils who had mastered a particular skill was greater in the groups which implemented the program for one year than in comparable groups which had not implemented the program. The positive effects of the preceding analysis were generally confirmed by results on subtests of the Doren Diagnostic Reading Test

administered to children in their third year. Mixed results were obtained on standardized tests of vocabulary and comprehension administered in the two pilot schools; in only one of the two pilot schools were consistently positive effects observed.

TABLE 1

DATA COLLECTION SCHEDULE FOR PILOT IMPLEMENTATION OF THE WISCONSIN DESIGN FOR READING SKILL DEVELOPMENT IN TWO SCHOOLS IN 1969 AND 1970

	School					
	A			B		
	Year in School			Year in School		
	2	3	4	2	3	4
Word Attack : Number of objectives mastered	Sept.	Sept.	Sept.			
Doren Diagnostic Reading Test		May			May	
Stanford Achievement Reading Subtests		Dec.	Mar.			
Gates-MacGinitie Reading Tests				May	May	May

TABLE 2

PERCENT OF BASELINE NONEXPERIENCED AND EXPERIENCED CHILDREN WHO MASTERED VARIOUS READING SKILLS

Level Administered	Group and year	N	Skill Number																
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
A	B-2 - 1969	134	48	63	79	94	65	--	30											
A	Ex-2 - 1970	107	62	52	79	97	66	--	37											
B	B-3 - 1969	105	26	--	74	43	58	--	32	6	40	13	21	55*	24					
B	Ex-3 - 1970	117	51	--	68	66	66	--	27	34	35	20	32	48*	58					
C	B-4 - 1969	113	67	17	82	63	23	46	31	--	--	--	--	58	46	76*	44	11	--	48
C	Ex-4 - 1970	100	83	35	85	76	58	63	51	--	--	--	--	70	65	56*	61	45	--	62

* Tests substantially changed, 1969 to 1970.

TABLE 3

DISTRIBUTION OF SKILLS MASTERED AND RETAINED BY THREE GROUPS

OF CHILDREN DURING 1969-1970 SCHOOL YEAR

Year in School in Sept. 1970	Number of Skills Mastered									Median[1]
	0-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	24---	
Third (N = 98)[2]	13	21	24	28	11	1				8
Fourth (N = 87)	3	4	4	3	10	13	22	1	27	19
Fifth (N = 96)	27	10	6	24	24	5				11

[1] Median is determined from raw, not grouped data.

[2] Numbers are smaller than in Table 3 because only those remaining in school one academic year and who were in school attendance during the week of testing could be included.

TABLE 4

MEAN RAW SCORES OF TWO SUCCESSIVE GROUPS OF CHILDREN COMPLETING

THEIR THIRD YEAR IN TWO SCHOOLS ON THE DOREN DIAGNOSTIC TEST IN MAY, 1969 AND MAY, 1970

Subtest	Baseline May 1969 N = 112	Experienced May 1970 N = 102	Difference	Baseline May 1969 N = 95	Experienced May 1970 N = 87	Difference
Letter Recognition (10)*	9.1	8.9	-.2	8.2	9.1	.9
Beginning Sounds (10)	8.8	8.8	0.0	8.4	8.7	.3
Word Recognition (15)	14.6	14.6	0.0	14.2	14.6	.4
Speech Consonants (5)	4.3	4.5	.2	4.4	4.5	.1
Ending Sounds (15)	11.3	11.5	.2	9.7	10.6	.9
Blending (10)	7.1	8.2	1.1	6.9	7.4	.5
Rhyming (10)	5.5	6.2	.7	4.9	5.3	.4
Vowels (25)	18.7	20.3	1.6	16.2	18.6	2.4
Homonyms (5)	3.9	4.0	.1	3.8	4.2	.4
Total Score (105)	83.4	86.5	3.1**	76.7	83.0	6.3***

* Indicates total number of items in the subtest

** p < .20

*** p < .01

TABLE 5

PERFORMANCE IN GRADE EQUIVALENTS OF CHILDREN IN SCHOOL A ON THE

READING SUBTESTS OF THE STANFORD ACHIEVEMENT TEST IN GRADE EQUIVALENTS *

Year in School	Administration Date	Word Study Skills			Paragraph Meaning			Vocabulary (Primary I) or Word Meaning (Primary II)		
		Q_1	Median	Q_3	Q_1	Median	Q_3	Q_1	Median	Q_3
3	December, 1969	1.9	2.2	2.8	1.7	2.0	2.6	2.1	2.5	3.1
	December, 1970	1.8	2.2	2.7	1.6	1.7	2.3	1.9	2.4	2.7
4	March, 1969	2.3	3.3	5.0	2.6	3.3	4.1	2.7	3.7	4.2
	March, 1970	2.4	3.3	5.0	2.3	3.1	4.0	2.7	3.5	4.0

* Children in their third year of school (Grade 2) were administered the Primary I test, those in their fourth year, the Primary II test.

TABLE 6

PERFORMANCE IN GRADE EQUIVALENTS OF CHILDREN IN SCHOOL B ON THE GATES-MAC GINITIE READING TESTS *

Year in School	Date of Test Administration	Vocabulary			Comprehension		
		Q_1	Median	Q_3	Q_1	Median	Q_3
Second	May, 1969	1.5	1.8	2.4	1.5	1.6	2.4
	May, 1970	1.7	2.2	2.8	1.6	1.8	2.5
Third	May, 1969	2.0	2.5	3.6	1.8	2.5	3.6
	May, 1970	2.0	3.0	4.3	1.8	2.6	4.5
Gain during third year		(0.5)	(1.2)	(1.9)	(0.3)	(1.0)	(2.1)
Fourth	May, 1969	3.0	3.5	4.4	2.7	3.7	4.6
	May, 1970	3.2	3.6	4.9	3.0	3.6	4.9
Gain during fourth year		(1.2)	(1.1)	(1.3)	(1.2)	(1.1)	(1.3)

* Children in their second year of school (Grade 1) were administered Form A, those in their third year, Form B, and those in their fourth year, Form C.