DOCUMENT RESUME

ED 028 126 SP 002 295

By-Wallen, Carl J.; And Others

Training Teachers to Determine Learner Achievement of Objectives in Reading Instruction.

Oregon Univ., Eugene. Coll. of Education.
Pub Date Feb 69
Note-14p.; Paper presented at the Annual Meeting of the American Educational Research Association in Los Angeles, February 1969.
EDRS Price MF-$0.25 HC-$0.80
Descriptors-Elementary School Teachers, *Methods Courses, Preservice Education, *Reading Diagnosis, Test Construction

The effectiveness of an undergraduate program in elementary school reading methods, designed to train the prospective teacher to determine learner achievement in word recognition, comprehension, and rate and study was investigated. Student teachers participated in simulated teaching experiences through self-correctional workbooks and, in the first two reading areas, through viewing video tapes of children reading. Each student teacher was then requested to write an informal test which would measure pupil attainment of a specific reading objective and fulfill three criteria: it would require a pupil response, would describe an adequate response, and would resemble the three reading categories as defined by the investigators. The same 20 tests were graded by four investigators to improve reliability. Each program part or reading area was administered in sequence and had four treatment groups: pretest-program-posttest (A), pretest-posttest (B), program-posttest (C), and posttest (D). It was found that treatments A and C were superior to B and D, but that A was not significantly different from C nor B from D. All three parts of the instructional program had a significant effect on the student teachers. (LP)

# TRAINING TEACHERS TO DETERMINE LEARNER ACHIEVEMENT OF OBJECTIVES IN READING INSTRUCTION*

Carl J. Wallen, Cyrus Blair, Douglas Robertson, Daniel Schooley

University of Oregon

This study investigated the effectiveness of the multimedia instructional program, DIAGNOSTIC TECHNIQUES IN READING FOR CLASSROOM TEACHERS. The instructional program was designed to train prospective elementary school teachers to determine learner attainment of objectives in three areas of reading: word recognition, comprehension, and rate and study.

The instructional program tested is the first phase of a two-phase program designed to train student teachers to use a Test-Teach-Test instructional strategy in teaching reading to elementary school children. The instructional strategy is adaptable to any "reading method."

The major instructional objective for the tested instructional program was:

Task. Given any objective in word recognition, comprehension, or rate and study which is stated in the unoperationalized manner commonly found in teacher's guides and manuals, e.g., "phonics--initial consonant b" or "know the meaning of stem," and asked to write an informal test which could be used to determine if a child had already attained the reading objective;



*Paper presented at the Annual Meeeing of the American Educational Research Association in Los Angeles, February, 1969.

Criteria. Student teachers will be able to write a test having three features: (1) a task requiring a pupil response, (2) the specific pupil response which will be considered adequate, and (3) a satisfactory resemblance between the task and adequate response suggested and the categories of word recognition, comprehension, and rate and study defined by the investigators.

The instructional program provided a series of simulated teaching experiences with children through the use of a self-correctional workbook correlated with videotaped episodes of children reading. The instructional program contained three parts, one for each area of reading ability: Part I, Word Recognition, Part II, Comprehenseion, and Part III, Rate and Study. The three parts were administered separately. Subject performance on each part was determined. The instructional program was self-correctional so all treatments were administered without the intervention of a college instructor.

## Part I, Word Recognition

Part I consisted of a 30-minute videotape and a 75-page self-correctional workbook. The total administration time ranged from four to six hours. Subjects spent the first hour in a television viewing room where they simulated the administration of informal tests for word recognition abilities by observing scenes of a teacher administering informal teacher-made tests. Subjects used their workbooks to record and later interpret the children's responses. The next three to six hours of treatment were accomplished in a study area selected by the subjects. Subjects completed the exercises in the workbooks: the exercises provided guided practice in administering simulated informal tests

and in designing informal tests for statements of reading abilities. The time to complete the workbook was not controlled because the organization of the workbooks was intended to allow for self-pacing. Subjects were asked to work alone. In discussions after the completion of the study, subjects claimed they had worked alone. To avoid a novelty effect, the program was administered as a regular part of an undergraduate reading methods course. Grouping was explained as being necessitated by the small size of the television viewing room.

Part II, Comprehension

Part II consisted of a 30-minute videotape and an 86-page self-correctional workbook. The administration times and conditions were identical to those described for Part I, except that the program content dealt with comprehension abilities.

Part III, Rate and Study

Part III consisted of a 45-page workbook. The simulated situations were based on written data presented in the workbook. The administration time was two to four hours. Again, subjects worked in a study area of their own choosing, at their own pace, and alone.

## METHOD

Criterion Measure

Six performance tests and scoring keys were constructed from the instructional objectives. Two performance tests were constructed for each part of the program: one pretest and one posttest. No attempt was made to determine the equivalence of pretests and posttests because the statistical analysis did not use performance data from the pretest. The

pretest was considered an additional treatment, whose effect must be assessed. The tests used an essay format. Subjects were given two un-operationalized objectives and asked to write an informal test for the objective.

A scoring key was constructed for each test. A ten-point scale was used in assessing the degree to which subjects had fully attained the criteria. The ten-point scale consisted of four levels of approximation between subject response and the categories defined by the investigators.

Level 1 (1 point). Subject described a situation which did not require a pupil response, and thus was not considered a test.

Level 2 (2 to 4 points). Subject gave a test, but the test could not be classified as having a satisfactory resemblance to the categories defined by the investigators. Points 3 and 4 reflected the completeness with which subjects described the anticipated pupil response.

Level 3 (5 to 7 points). Subject gave a test which could be classfied as having a satisfactory resemblance to the categories defined by the investigators, but the test was inadequate because it lacked specificity or sufficient length. Points 6 and 7 reflected the completeness with which subjects described the anticipated pupil response.

Level 4 (8 to 10 points). Subject gave a test which could be classified as having a satisfactory resemblance to the categories defined by the investigators and had a satisfactory specificity and length. Points 9 and 10 reflected the completeness with which subjects described the anticipated pupil response.

The reliability of the scoring procedure was measured by having the four investigators separately score the same 20 tests. The tests had

the same format as those used in the study. The tests had been administered the previous term to a different group of students enrolled in a similar undergraduate reading methods course.

An analysis of variance to estimate reliability was performed on the scores obtained, the results of which are reported in Table 1. The reliability coefficient for the four investigators was .975.

Table 1

Analysis of Variance Summary Table

| Source | S.S. | df | M.S. | F | P |
|---|---|---|---|---|---|
| Ratees (rows) | 620.8 | 19 | 32.67 | 40.403 | .001 |
| Raters (col.) | 2.9 | 3 | .97 | 1.196 | (N.S.) |
| Remainder | 46.1 | 57 | .81 | | |
| Total | 669.8 | 79 | | | |

Two procedures were used to assure investigator impartiality in scoring tests. (1) Investigators scored a random assortment of papers without knowing to which of four possible treatments each test belonged. (2) Each investigator scored only one of the two sections that made up each subject's test, and did not know how the subject had performed on the other half of the test.

Treatments

A four-group design was used (Campbell & Stanley, 1968). The four-group design allowed the determination of the main effects of the instructional program and of the pretests. The organization of the four treatments in the design is shown in Illustration 1.

Illustration 1

The Four-Group Design

| | | | | |
|---|---|---|---|---|
| R | (Treatment A) | $O_1$ | X | $O_2$ |
| R | (Treatment B) | $O_1$ | | $O_2$ |
| R | (Treatment C) | | X | $O_2$ |
| R | (Treatment D) | | | $O_2$ |

Where:
O indicates a performance test. (Test 1 and 2 are indicated.)
X indicates administration of instructional program.
R indicates random selection of subjects.

The four-group design was repeated for each part of the instructional program: Part I, Part II, Part III. The three parts of the instructional program were designed to be administered in sequence. The repeated design allowed the effectiveness of each part to be determined within the context of the total instructional program. In order to prepare all subjects for Parts II and III additional post-experiment treatment sessions were arranged after the administration of Parts I and II. The repeated design with post-experiment treatments is shown in Illustration 2.

Illustration 2

The Four-Group Design Repeated Three Times
with Post-Experiment Treatments

| Treatment. | Part I | (N) | Part II | (N) | Part III | (N) |
|---|---|---|---|---|---|---|
| A | R-$O_1$ $X_I$ $O_2$- | 13 | R-$O_3$ $X_{II}$ $O_4$- | 16 | R-$O_5$ $X_{III}$ $O_6$- | 13 |
| B | R-$O_1$ $O_2$- $X_I$ | 14 | R-$O_3$ $O_4$- $X_{II}$ | 13 | R-$O_5$ $O_6$- $X_{III}$ | 16 |
| C | R- $X_I$ $O_2$- | 11 | R- $X_{II}$ $O_4$- | 14 | R- $X_{III}$ $O_6$- | 12 |
| D | R- $O_2$- $X_I$ | 16 | R- $O_4$ $X_{II}$ | 13 | R- $O_6$- $X_{III}$ | 12 |
| Total N | | 54 | | 57 | | 53 |

Where:
O indicates a performance test. (Six were administered.)
X indicates administration of a part of the instructional program: I, II, III.
R indicates random selection of subjects.

Subjects were enrolled in two of the three sections of the undergraduate course in elementary school reading methods at the University of Oregon during the fall term, 1968. There was no reason to suspect that subjects in the two participating sections differed from those in the one non-participating section as the original assignment to sections had been random.

Each participating section had the four treatments. Subjects were randomly assigned by section. The total number of subjects in each treatment was not equal due to absences. When a subject assigned to Treatment A or Treatment C missed his scheduled instructional program, his performance test score was not considered in the data analysis for that part of the instructional program.

Those subjects were given that part of the instructional program during the post-experiment treatment sessions arranged for Treatments B and C.

## RESULTS

The investigators used three distinct, one-way analyses of variance with the four treatment levels for each of the three parts of the instructional program. Since the number of subjects in each of the treatments was unequal, a weighted solution was computed (Winer, 1962). The results for the three analyses are reported in Tables 2, 3, and 4.

Table 2

Analysis of Variance Summary Table for Part I

| Source | SS | df | M.S. | F |
|---|---|---|---|---|
| Treatments | 593.92 | 3 | 197.97 | 25.95* |
| Experimental error | 381.58 | 50 | 7.63 | |
| Total | 975.50 | 53 | | |

* $p < .001$

Table 3

Analysis of Variance Summary Table for Part II

| Source | SS | df | M.S. | F |
|---|---|---|---|---|
| Treatments | 663.29 | 3 | 221.10 | 17.26* |
| Experimental error | 666.07 | 52 | 12.81 | |
| Total | 1,329.36 | 55 | | |

* $p < .001$

Table 4

Analysis of Variance Summary Table for Part III

| Source | SS | df | M.S. | F |
|---|---|---|---|---|
| Treatments | 1,014.68 | 3 | 338.23 | 22.08* |
| Experimental error | 750.53 | 49 | 15.32 | |
| Total | 1,765.21 | 52 | | |

* $p < .001$

The treatment factors were significant ($p < .001$) for each of the three parts of the instructional program. The significance indicated differential effects among the treatments.

The Newman-Keuls method (Winer, 1962) was adapted to test the difference between all pairs of means for each of the parts of the instructional program. The results of the Newman-Keuls analyses for each of the three parts are reported in Tables 5, 6, and 7.

Table 5

Tests on Differences Between All Pairs of Means

in Part I.

| Order | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Treatments in Order of $T_j$ | D | B | C | A |
| $T_j$ ..... | 3.69 | 4.86 | 10.45 | 11.15 |

| | D | B | C | A |
|---|---|---|---|---|
| D | -- | 1.17 | 6.76 | 7.46 |
| B | -- | -- | 5.59 | 6.29 |
| C | -- | -- | -- | .78 |
| A | -- | -- | -- | -- |

| Truncated range r | 2 | 3 | 4 |
|---|---|---|---|
| $q.99^{(r,50)}$ | 3.79 | 4.33 | 4.65 |
| $\frac{q.99^{(r,50)}}{\sqrt{nMS_{error}}}$ | 2.88 | 3.28 | 3.53 |

| | D | B | C | A |
|---|---|---|---|---|
| D | | | ** | ** |
| B | | | ** | ** |
| C | | | | |
| A | | | | |

Table 6

Tests on Differences Between All Pairs of Means

in Part II

| Order | | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Treatments in order of $T_j$ | | D | B | C | A |
| $T_j$ | | 5.38 | 6.07 | 12.38 | 12.77 |
| | D | B | C | A | |
| D -- | | .69 | 7.00 | 7.39 | |
| B -- | | -- | 6.31 | 6.70 | |
| C -- | | -- | -- | .39 | |
| A -- | | -- | -- | -- | |
| Truncated range r | | 2 | 3 | 4 | |
| $q_{.99}^{(r,52)}$ | | 3.78 | 4.32 | 4.64 | |
| $q_{.99}^{(r,52)}/\sqrt{nMS_{error}}$ | | 3.63 | 4.15 | 4.45 | |

| | D | B | C | A |
|---|---|---|---|---|
| D | | | ** | ** |
| B | | | ** | ** |
| C | | | | -- |
| A | | | | |

Table 7

Tests on Differences Between All Pairs of Means

in Part III

| Order | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Treatments in order of $T_j$ | D | B | C | A |
| $T_j$ | 5.58 | 7.19 | 13.92 | 16.25 |

| | D | B | C | A |
|---|---|---|---|---|
| D | -- | 1.61 | 8.34 | 10.67 |
| B | -- | -- | 6.73 | 9.06 |
| C | -- | -- | -- | 2.33 |
| A | -- | -- | -- | -- |

| Truncated range r | 2 | 3 | 4 |
|---|---|---|---|
| $q.99^{(r,49)}$ | 3.79 | 4.33 | 4.66 |
| $q.99^{(r,49)}\sqrt{nMS_{error}}$ | 4.10 | 4.69 | 5.04 |

| | D | B | C | A |
|---|---|---|---|---|
| D | | | ** | ** |
| B | | | ** | ** |
| C | | | | |
| A | | | | |

The results of the Newman-Keuls method indicated that Treatments A and C were superior to Treatments B and D ($p < .01$), but Treatments A and C were not significantly different from each other, nor did Treatments B and D differ significantly from each other. All three parts of the instructional program had a significant effect. The pretests administered for each part of the instructional program did not have a significant effect.

## SUMMARY AND CONCLUSIONS

This study investigated the effectiveness of a multimedia instructional program designed to train prospective elementary school teachers to determine learner attainment of objectives in three areas of reading: word recognition, comprehension, and rate and study. The instructional program contained three parts, one for each area of reading. The effectiveness of each part was separately determined.

Performance measures were constructed from the instructional objectives for each part of the instructional program. A ten-point scale was used in assessing the deg to which subject responses approximated the defined criteria. The reliability of the scoring procedure was determined by having the four investigators separately score the same 20 tests. The reliability was .975.

A four-group design was used in order to determine the effect of each part of the instructional program and the effect of the pretests. Treatment A had pretest-program-posttest. Treatment B had pretest-posttest. Treatment C had program-posttest. Treatment D had posttest.

Three distinct, one-way analyses of variance were used with the four treatment levels for each of the three parts of the instructional program. The treatment factors were found to be significant ($p < .001$) for each part of the instructional program.

The Newman-Keuls method was adapted to test the difference between all pairs of means for each of the three parts of the instructional program. The results indicated that Treatments A and C were superior to Treatments B and D ($p < .01$), but Treatments A and C were not significantly different from each other, nor did treatments B and D differ significantly from each other. All three parts of the instructional had a significant effect. The pretests administered for each part of the instructional program did not have a significant effect.

The multimedia instructional program DIAGNOSTIC TECHNIQUES IN READING FOR CLASSROOM TEACHERS was effective with undergraduate students enrolled in a reading methods course at the University of Oregon. The instructional program would probably be effective with undergraduate students enrolled in reading methods courses at other teacher training institutions.

The instructional program trains prospective elementary school teachers to determine learner attainment of objectives in three areas of reading: word recognition, comprehension, and rate and study. A teacher who can determine learner attainment of specific objectives in reading will be able to provide that particular instructional exercise which is most appropriate for any one child.

## REFERENCES

Campbell, D. T., and Stanley, J. C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally and Company, 1968, pp.24-25.

Winter, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Company, 1962, p. 124.