DOCUMENT RESUME

ED 127 945 IR 003 909

AUTHOR Hodapp, Timothy
TITLE READALONG in Standard Classes, Evaluation Report.
INSTITUTION Ontario Educational Communications Authority, Toronto. Research and Planning Branch.
REPORT NO OECA-PRBR-3B
PUB DATE May 76
NOTE 19p.; For a related document see IR 003 908

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Comparative Analysis; Conventional Instruction; *Early Childhood Education; Grade 1; Grade 2; *Instructional Aids; Kindergarten; *Media Research; Program Evaluation; Reading Instruction; *Video Tape Recordings; *Vocabulary Skills
IDENTIFIERS *Readalong

ABSTRACT

"Readalong," a program employing a videotape format to teach vocabulary skills to early childhood learners, was tested in a standard classroom setting by comparing kindergarten, first, and second grade classes using the program with classes using only the conventional materials. At three points during the ten-week series, students were asked to read a list of program and non-program words. The program proved to be effective, and it was recommended that the format and techniques not be altered. This report discusses research methodology and summarizes results. (EMH)

The Ontario Educational Communications Authority

Research and Planning Division

# READALONG

# IN STANDARD CLASSES

# EVALUATION REPORT

Project Manager: K. G. O'Bryan, Ph.D

Project Officer: Timothy Hodapp

Research Assistance from: Sr. Patricia Donovan
Deborah Cheeseman

Secretary: Sally K. Millar

May, 1976

**The Ontario Educational Communications Authority**

**Research and Planning Division**

ABSTRACT A - 3B

# READALONG IN STANDARD CLASSES

## EVALUATION REPORT

By: Timothy Hodapp

(OECA Project Research Branch Report Number 3B, 7pp, May 1976)

One of the main objectives of Readalong is to provide primary readers with a basic vocabulary which they can use as a foundation for building other reading skills. This evaluation was designed to determine if the program was attaining this goal.

Several teachers of senior kindergarten, Grade 1, and Grade 2 classes who indicated they were using the program volunteered to assist in the evaluation. At three points during the 10 week series their students were asked to read a list of program and non-program words. At the same times control students from the same grades were asked to read the same lists. Difference scores between program words read and non-program words read were calculated. Several special education students were also tested, but without a control group.

Results indicated that the experimental Grade 1 group had significantly higher positive difference scores throughout the series. The experimental senior kindergarten class had a significantly higher positive difference score only for the first testing period, and the experimental Grade 2 group only for the middle testing period. The special education class knew more program words than non-program words only during the first period. Analysis also indicated that performance on the program words decreased over the course of the program in relation to the non-program words.

Due to the effectiveness of the program with its primary audience it was recommended that the format and techniques not be altered. However, because of the overall decreasing performance on program words it was recommended that programs be scheduled further apart or that teachers be encouraged to maintain consistent utilization of the program.

Senior kindergarten and special education teachers should also be encouraged to use the program at the appropriate times of the school year when their students can handle it and, perhaps, with specially developed resource material.

READALONG IN STANDARD CLASSES

EVALUATION REPORT

## I. PURPOSE

This project was initiated as part of a continuing effort to provide the producers and educators of Readalong with feedback on the effectiveness of the program in aiding reading instruction to primary and special education students. This evaluation extends the previous evaluation (Readalong Pilot Programs, O.E.C.A. Project Research Report #3, 1975) in that: all thirty programs of the series were employed in the evaluation, the program was used by teachers in regular classroom settings as part of their regular classroom instructions, the teachers had preliminary copies of the teacher's guide for the program, students from a wider range of grade levels were used as subjects, and a control group was used for purposes of comparison.

The educators for Readalong have applied (as part of a basic curriculum for primary readers) a widely held belief among reading researchers that a basic vocabulary of 30 - 50 words is required by new readers before they can be expected to progress effectively to the more difficult aspects of reading. Thus, although all words were taught within a context of conversational English, this evaluation was limited to the extent that it tested the ability of students to read isolated program words in comparison to isolated non-program words after exposure to the program.

Other, more specific, objectives have also been defined by the educators of the series. These include such goals as improving or developing students' desire to read and their confidence in solving reading related problems. The effectiveness of Readalong in achieving these objectives will be discussed in subsequent reports.

## II. METHODOLOGY

### Subjects

Several teachers from the North York School Board indicated that they were using the program and would participate in the evaluation. Their entire classes were used as the experimental group.

The control group was formed from the students of another North York school. All schools were situated in essentially middle class neighborhoods. All classes were composed of average students with an average range of abilities.

A total of 330 students were tested during the course of the evaluation. However, due to attrition, 189 comprised the final sample. Sample sizes for each of the grades of the experimental group were: Senior Kindergarten - 2 classes, 27 students; Grade 1 - 4 classes, 54 students; and Grade 2 - 1 class, 10 students. The arrangement of subjects in the control group was: Senior Kindergarten - 2 classes, 34 students; Grade 1 - 3 classes, 45 students; and Grade 2 - 2 classes, 19 students.

In addition, 3 students from a remedial class from a North York school were tested. Due to the difficulty in arranging for a properly matched control group, one was not used.

### Procedure

The Readalong series comprises 30 programs with three being aired each week. The first program of the week is broadcast on Monday and Tuesday, the second on Wednesday and Thursday, and the third on Friday. Each program is broadcast twice a day, once in the morning and once in the afternoon. Classes in the experimental group viewed the programs as a group in their individual classrooms or in the school libraries.

Teachers were given preliminary copies of the guidance material. While this package was incomplete, it did include the essential elements: the guidelines for classroom use and activity sheets describing the content of each program and suggestions for related activities.

No restrictions can be imposed on the use of Readalong or the guidance materials when the program is in normal use. Therefore, no restrictions were suggested or imposed on the teachers of the experimental group classes in their use of the program or ancillary materials.

Subjects were tested individually in a location away from noise or distraction, usually a quiet corner of the library or the staff room. Subjects were presented with a list of 24 words, printed individually on flashcards. Twelve were program words and 12 were non-program words.

The order of the 24 words was randomized for each student. Each subject was requested to read the words if they could or to indicate non-comprehension if they could not. If the child could not read a word after 10 seconds, the experimenter proceeded to show the next word. After the testing session the child was returned to his or her classroom. This method was employed in order to minimize disruption of normal classroom activities as much as possible.

Word Lists.

Subjects were tested three times. At each testing session a completely different set of words was employed. The twelve program words were randomly selected from the words introduced in the Readalong programs since the last testing session.

The words used in the three non-program word lists were randomly selected from the Level I Gaites-MacGinitie Test, a test designed to determine the reading ability of primary level students. Thus, the non-program word lists are representative of words within the grasp of the average Grade 1 student.

Since the control words were randomly selected from a single list, it can be assumed that the lists do not significantly differ in difficulty from one another. However, because the program words were chosen arbitrarily, it cannot be assumed that these lists do not significantly differ in difficulty from one another or from any list of control words. All word lists are provided in Appendix A.

## III. RESULTS

Normal Classes.

Each student received three scores for each of the testing sessions: the number of program words read correctly, the number of novel words read correctly, and the difference between the number of program words and novel words read correctly. Analysis of the difference scores will provide the greatest information concerning the effectiveness of the program in that this measure controls for reading ability; i.e.; readers unaffected by the program should read both program and control words with the same facility, whether with ease or difficulty; and it controls for different levels of difficulty between the various word lists, should such difference exist, i.e., if the program is ineffective but word lists are of different difficulty levels, the difference scores between control and experimental subjects should be equal.

Data was analyzed using the repeated measures analysis of variance program from the Biomedical Computer Programs package and appropriate procedures from the Statistical Package for the Social Sciences. Significance level for rejection of null hypotheses was $p = .01$.

Figure I provides a graphic display of the mean difference scores for each grade in each condition. See Tables 1 - 3 in Appendix B. A positive score indicates more program words read than non-program words. As can be seen at Test Period I (T1) all of the experimental grades had higher positive differences than their control counterparts. However, while the experimental grades and the control grades had similar difference scores among themselves (EXP: $F(2,88) = .51$, NS; CON: $F(2,88) = 2.33$, NS) and the experimental grades as a whole performed better than the control grades as a whole ($F(1,183) = 17.61$, $P < .001$) significant differences in difference scores between conditions occurred only for the Senior Kindergarten and Grade 1 groups (SK: $t = 5.21$, $p < .001$; Gr.1: $t = 5.90$, $p < .0001$; Gr.2: $t = 1.74$, NS)

At Test Period 2 (T2) difference scores for all groups have decreased, with the score for the experimental kindergarten class having dropped dramatically. The differences between grades for the experimental group are now significant ($F(2,88) = 8.14$, $p < .001$) with significant differences occuring between the Kindergarten group and the Grade 1 group ($t = 3.96$, $p < .0001$) and the combined Grade 1 and Grade 2 groups ($t = 3.52$, $p < .001$). The differences between the grades of the control group are non-significant ($F(1,183) = 1.04$). However, further analysis indicates that significant differences occur between the experimental and control conditions for the Grade 1 and Grade 2 groups (SK: $t = 2.07$, NS; Grade 1: $t = 7.08$, $p > .0001$; Grade 2: $t = 3.86$, $p < .001$).

The grade main effect is significant ($F(2,183) = 11.20$, $p < .001$) as can be seen in that difference scores increase as grade increases. The grade by conditions interaction is also significant ($F(2,183) = 9.54$, $p < .001$) in that the difference scores increase over grade for the experimental group while they drop over grade for the control group.

At Test Period 3 (T3) the difference scores have again dropped in most cases, with the drop for the experimental Grade 2 group being most dramatic. The differences between grades are significant for both experimental and control conditions (EXP: $F(2,88) = 8.34$, $p < .001$; CON: $F(2,95) = 14.70$ $p < .001$). In the experimental group the Grade 1 group had a significantly higher difference score than both the Kindergarten group ($t = 2.78$, $p < .01$) and the Grade 2 group ($t = 3.54$, $p < .001$), both of which had negative difference scores. In the control group the Grade 2 group had a significantly lower difference score than the Kindergarten group ($t = 5.42$, $p < .0001$) and the Grade 1 group ($t = 3.68$, $p < .0001$). All grades in the control group had negative difference scores.

Overall, the condition main effect was not significant ($F(1,183) = .20$, NS). However, the experimental Grade 1 group had a significantly higher difference score than the control Grade 1 group ($t = 6.16$, $p < .0001$). The other two comparisons were not significant (SK: $t = .76$, NS; Gr.2: $t = 2.08$, NS). The grade main effect is significant ($F(2,183) = 9.63$, $p < .001$) as can be seen in that difference scores increase, generally, by grade. The grade by condition interaction is significant ($F(2,183) = 9.74$, $p < .001$) as can be seen in that the difference scores for the control group decrease faster by grade than the experimental grade difference scores.

FIGURE I

TEST OF DIFFERENCES

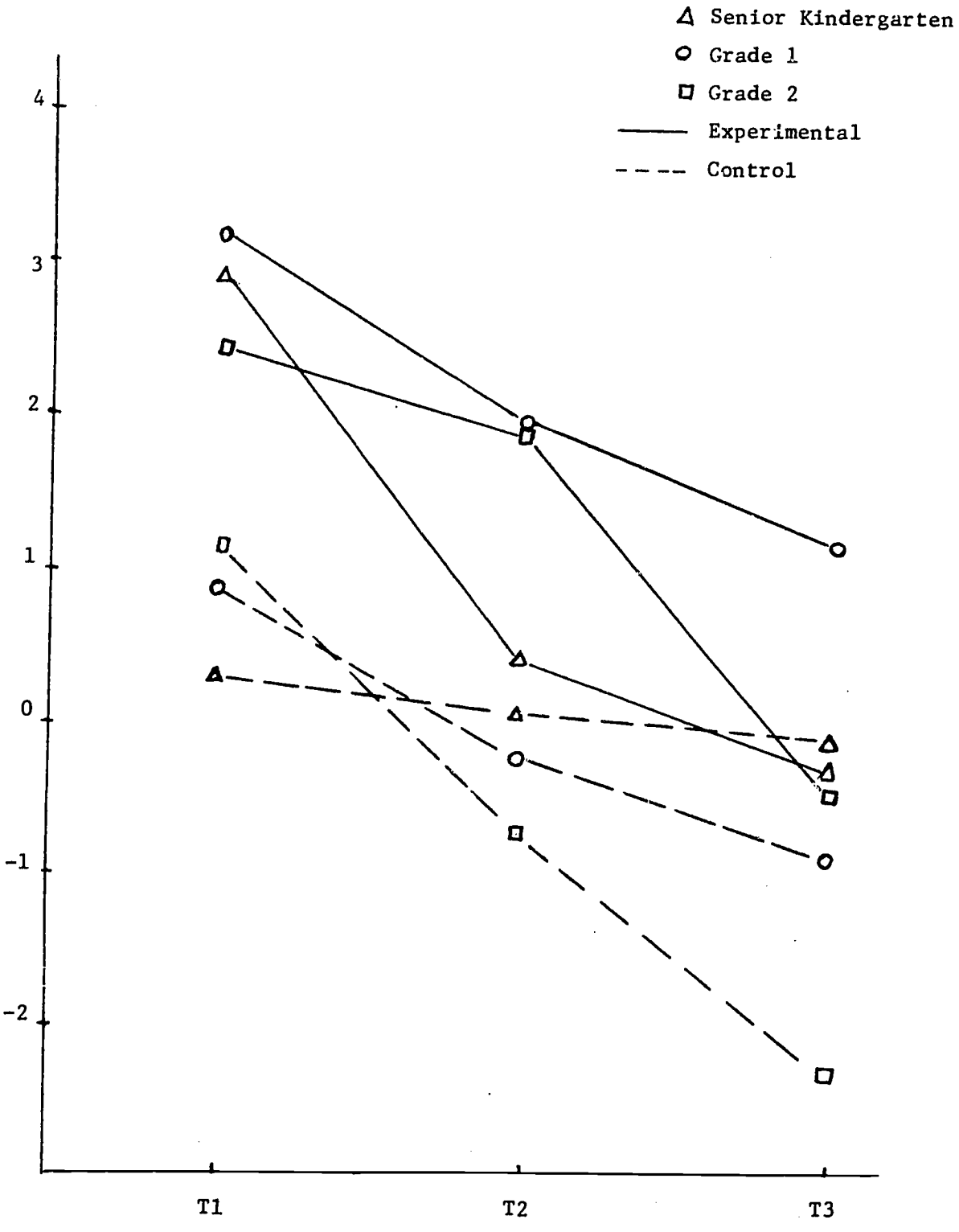


Figure 2 presents a graphic display of the mean difference scores between the entire experimental and control groups, i.e., grade was not considered. See Table 4 in Appendix B. A positive score indicates more program words read than non-program words. It can easily be seen that the experimental group had significantly higher difference scores over the course of the program ($F(1,183) = 50.43$, $p < .0001$). Further analysis also indicated that the differences between experimental and control groups were significant at each testing session (T.: $t = 8.36$, $p < .0001$; T2: $t = 7.94$, $p < .0001$; T3: $t = 5.70$, $p < .001$). However, it should be recalled that this is an overall analysis, i.e., significant differences between control and experimental groups were not exhibited at each testing session.

The difference scores also significantly decrease over the course of the program ($F(2,366) = 71.82$, $p < .0001$). However, the difference scores appear to decrease by approximately the same increment for both groups. The condition by time interaction and the grade by time interaction (not illustrated here) both proved to be non-significant (C x T: $F(2,366) = 2.58$, NS; G x T: $F(4,366) = 2.83$, NS).

These results indicate that either the program words or the control words became easier or harder over the course of the program. Analysis revealed that for the non-program words the main effect for condition was not significant ($F(1,183) = 1.78$, NS), nor was the time by condition interaction ($F(2,366) = 3.75$, NS). For the program words, the conditions main effect was significant ($F(1,183) = 21.33$, $p < .0001$) as was the time by condition interaction ($F(2,366) = 10.35$, $p < .001$). Thus, it appears that the decreasing difference scores were due to the changes in the program word scores in comparison to the relatively stable non-program word scores.

Special Education.

Figure 3 presents a graphic display of the mean scores for program and non-program words for the 3 special education students. See Table 5, Appendix B. Analysis revealed that the differences between the two were significant only at Testing Period I. (T1: $t = 23,00$, $p < .01$; T2: $t = 3.88$, NS; T3: $t = .75$, NS)

FIGURE 2

TEST OF DIFFERENCES

ALL GRADES

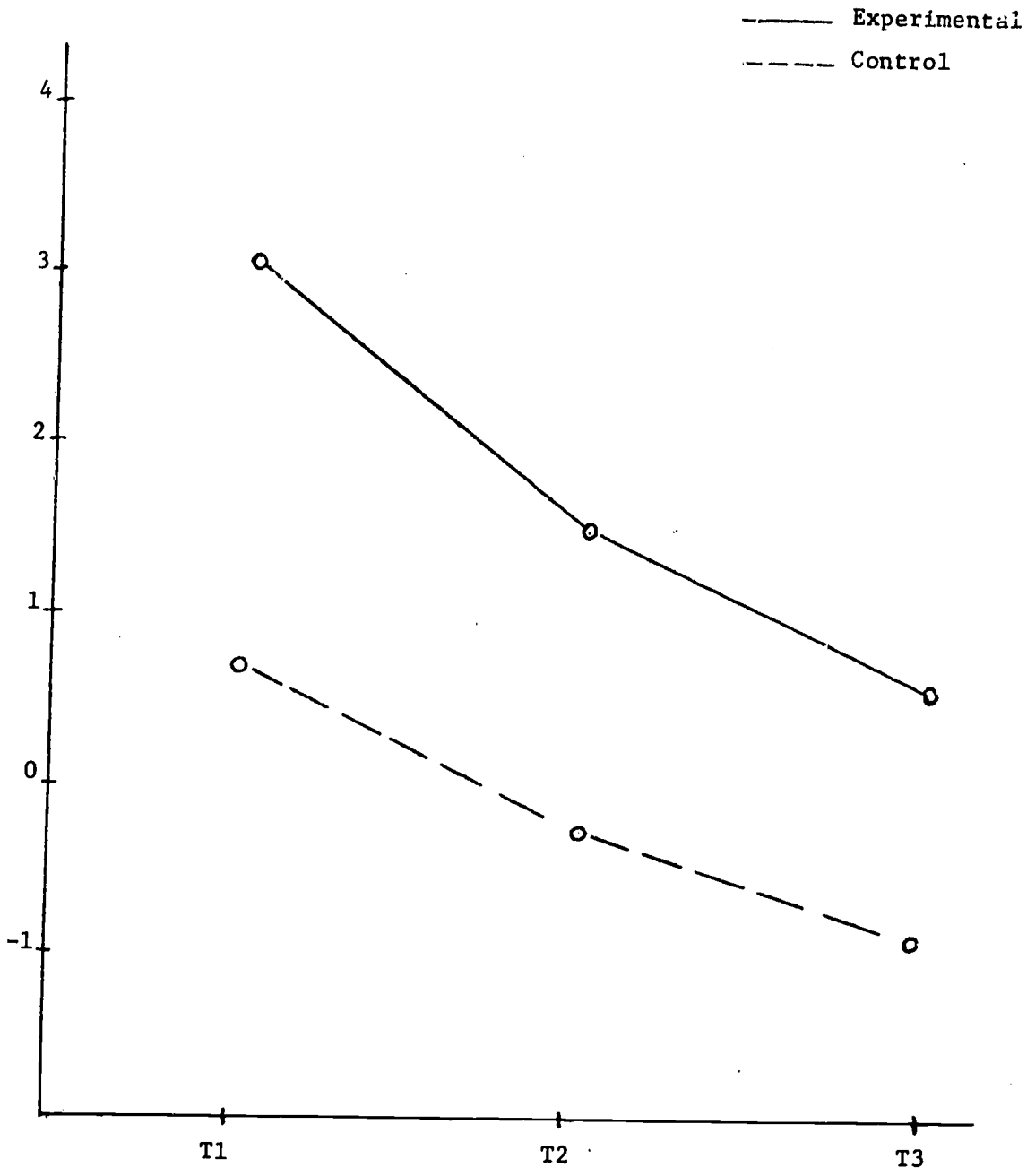
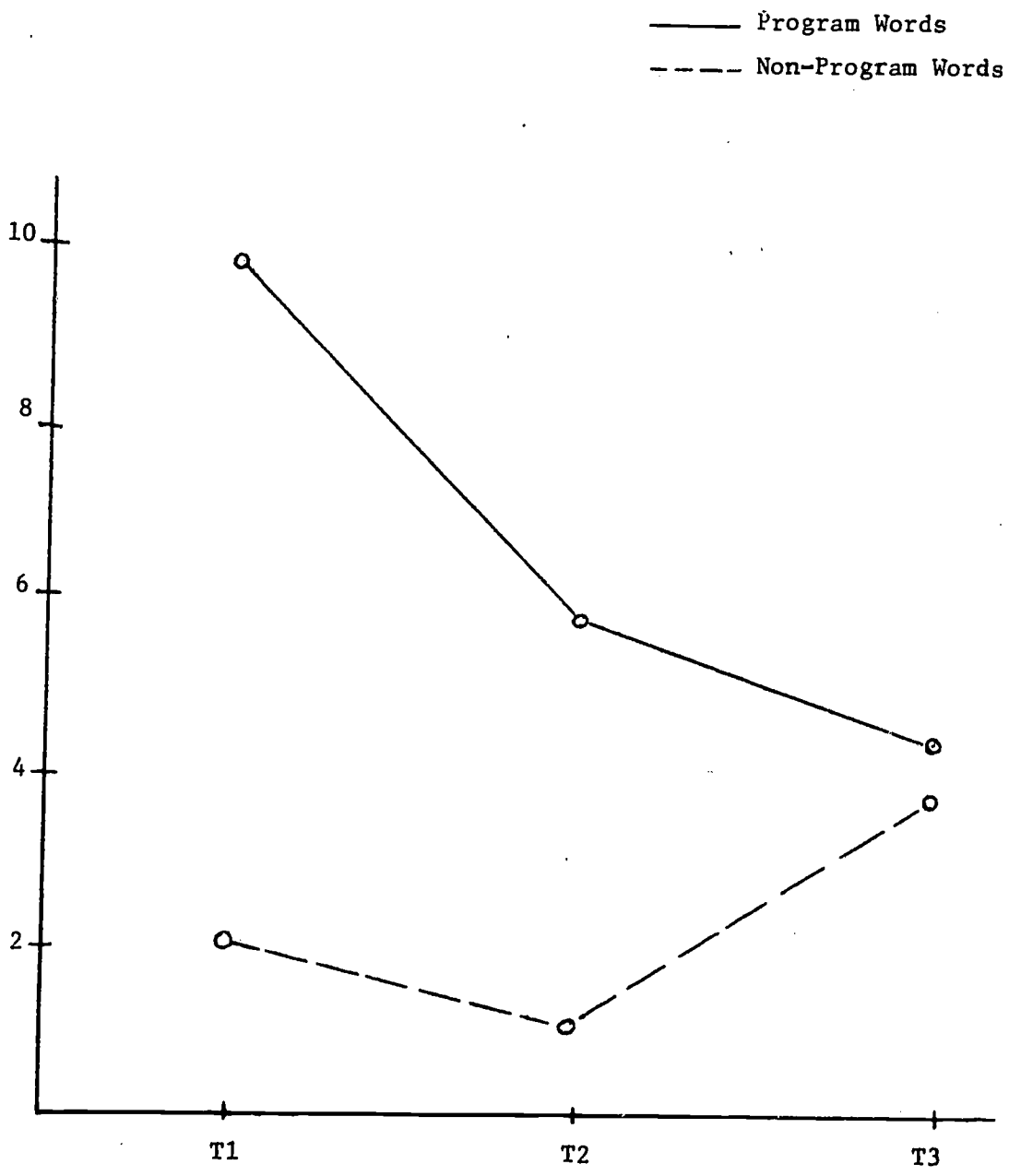


FIGURE 3

SPECIAL EDUCATION CLASS



## IV. DISCUSSION

Based on the results of the previous chapter it can be concluded that Readalong and its supporting materials are an effective educational system for the teaching of reading. However, the results indicate that the program was only fully effective for its primary audience, Grade 1 students. Only the Grade 1 students who viewed the program had consistently higher difference scores between program and non-program words than the control Grade 1 students.

The program proved effective for the experimental senior kindergarten group only for the first testing period. However, the first few weeks of the program are very basic, in which only simple words are introduced. After this, the pace quickens. Thus, it is not surprising that their performance dropped off.

A similar pattern emerged for the special education class. Their performance also dropped off after the first testing period, possibly for the same reasons.

The program proved effective for the Grade 2 group only during the middle testing period. It was expected that the words taught on the program would be too easy for the regular Grade 2 class, and, thus, no significant differences would emerge at all. It may be that since the program words were chosen arbitrarily, those chosen for the middle weeks were difficult for the average Grade 2 student.

It was also found that difference scores dropped over the course of the program, and that this seemed to be due to decreasing abilities to read the program words in relation to ability to read the non-program words. Several reasons for this finding are possible. The program words may increase in difficulty. The way the words were presented may have been less effective in later programs. This would be especially important if the words actually do become more difficult.

Finally, the teachers may not have utilized the program in a consistent manner throughout the series. During the first few weeks, the program is new and exciting for the teacher as well as the student. However, after the novelty has worn off the program may become another teaching aid in competition with other devices and activities.

## V. RECOMMENDATIONS

Based on the results and conclusions of this study, the following recommendations can be proffered:

1. Generally, the program format and techniques have proved successful and should be continued.

2. The pace of the program should be decreased to some extent in order to counteract the decreasing comprehension of program words. This could be accomplished by slowing the pacing of the program to one or two programs a week or by the inclusion of more review programs.

3. Teachers should be impressed with the necessity for maintaining a consistent integration of Readalong into their reading instruction. The development of a complete reading curriculum with Readalong as its foundation may be the best means for attaining this end.

4. Senior kindergarten teachers should be encouraged to use the program during the last portion of the school year.

5. Special education teachers should be encouraged to use the program, perhaps with materials developed especially for their situation. Later programs might be spaced or more review programs added to increase learning of program words.

6. When programs are made for Grade 2 classes, writers should look to the middle programs for the types of words to be used.

7. If a program is ever designed for developing reading readiness, writers should look at the first weeks of Readalong for the types of words to be used.

## WORD LISTS

### TEST PERIOD 1

#### Program Words

Boy
Room
Book
Push
House
Ball
Kick
Read
Play
Here
Rain
There

#### Non-Program Words

Egg
Work
Fire
Hard
Tree
Pig
Splash
Stand
Grass
Grow
Ride
Stick

### TEST PERIOD 2

#### Program Words

Swim
Hole
Thing
Why
Sad
How
Would
Quick
When
Safe
Dress
Front

#### Non-Program Words

Soup
Four
Made
Had
Puppy
Oven
Bell
Flag
Warm
Hall
Hose
River

### TEST PERIOD 3

#### Program Words

Broken
Elephants
True
Large
Small
Flowers
Great
Teacher
Race
Story
Noise
Shake

#### Non-Program Words

Again
Eye
Wolf
Sheet
Bird
Not
Draw
Pair
Help
Wood
Star
Prince

TABLE I

MEANS OF DIFFERENCES

TESTING PERIOD 2

| | GRADE | | |
|---|---|---|---|
| | SK | 1 | 2 |
| EXP | 2.96 | 3.17 | 2.40 |
| CON | 0.26 | 0.84 | 1.11 |

TABLE 2

MEANS OF DIFFERENCES

TESTING PERIOD 2

| | GRADE | | |
|---|---|---|---|
| | SK | 1 | 2 |
| EXP | 0.41 | 1.98 | 1.90 |
| CON | 0.03 | -0.24 | -0.79 |

TABLE 3

MEANS OF DIFFERENCES

TESTING PERIOD 3

| | GRADE | | |
|---|---|---|---|
| | SK | 1 | 2 |
| EXP | -0.30 | 1.11 | -0.50 |
| CON | -0.12 | -0.91 | -2.37 |

TABLE 4

MEANS OF DIFFERENCES

ALL GRADES

| | T1 | T2 | T3 |
|---|---|---|---|
| EXP | 3.02 | 1.50 | 0.52 |
| CON | 0.69 | -0.26 | -0.92 |

TABLE 5

MEANS

SPECIAL EDUCATION

| | T1 | T2 | T3 |
|---|---|---|---|
| Program Words | 9.67 | 5.67 | 4.33 |
| Non-Program Words | 2.00 | 1.00 | 3.67 |