DOCUMENT RESUME

ED 058 002 24 RE 003 943

AUTHOR	Thomas, Charles R.
TITLE	The Effects of Visual Tracking Training Upon Five Selected Aspects of the Silent Reading Performance of First Grade Pupils as Measured by Eye-Movement Photography. Final Report.
INSTITUTION	Maine Univ., Orono.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-9-A-035
PUB DATE	Oct 71
GRANT	OEG-1-9-090035-0110(010)
NOTE	173p.
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	*Eye Movements; *Grade 1; Sex Differences; *Silent Reading; Test Reliability; Transfer of Training; *Visual Perception

ABSTRACT

One hundred and fifteen first graders were randomly assigned to experimental and control groups. Experimental pupils used the Visual Tracking program, and the control pupils participated in directed listening activities in separate rooms. The teachers followed a weekly rotating schedule in supervising the groups. After 12 weeks of training, eye-movement of all pupils were photographed with the EDL Reading Eye Camera as they read silently. Forty-four pupils were retested to collect data for determination of test-retest reliability. Eye-movement scores of 96 pupils were analyzed by factorial analysis of variance. Results showed the following: (1) No significant differences were found between the two groups on any of the five eye-movement measures (fixations, regressions, span, duration, and rate). (2) Girls read with significantly fewer fixations and at a significantly faster rate than did the boys ($p < .05$). (3) No significant interaction effects were found between visual tracking training and sex. (4) Test-retest reliabilities for the five eye-movement measures ranged from .66 to .88, and (5) Inter-rater reliability coefficients ranged from .82 to .98. It was concluded that the proposition that visual tracking training would improve first graders' silent reading performance was not supported. Tables, appendixes, and a bibliography are included. (AW)

FINAL REPORT

Project No. 9-A-035
Grant No. OEG-1-9-090035-0110 (010) OE-BR

THE EFFECTS OF VISUAL TRACKING TRAINING UPON FIVE SELECTED ASPECTS OF THE SILENT READING PERFORMANCE OF FIRST GRADE PUPILS AS MEASURED BY EYE-MOVEMENT PHOTOGRAPHY

Charles R. Thomas
University of Maine
Orono, Maine 04473

October 1971

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

FINAL REPORT

Project No. 9-A-035
Grant No. OEG-1-9-090035-0110 (010)

THE EFFECTS OF VISUAL TRACKING TRAINING UPON FIVE SELECTED ASPECTS OF THE SILENT READING PERFORMANCE OF FIRST GRADE PUPILS AS MEASURED BY EYE-MOVEMENT PHOTOGRAPHY

Charles R. Thomas

University of Maine
Orono, Maine 04473

October 1971

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

ACKNOWLEDGMENTS

It is with sincere appreciation that I acknowledge the many persons who have contributed to this study. I could not fail to thank the staff members and pupils of Weatherbee Elementary School in Hampden, Maine who so patiently endured my intrusion into their regular school program. It was the wonderful cooperation of Mr. Willard Hillier, principal, and his first grade teachers, Mrs. Gertrude Mayhew, Mrs. Susan McPike, Mrs. Esther Patten, Mrs. Roxine Patterson, and Mrs. Nancy Sawyer, that made this study possible.

The excellent cooperation and assistance of Mr. Linwood White, principal, and the first grade teachers of Asa C. Adams School in Orono, Maine contributed significantly to the success of the pilot study.

Finally, I am grateful to Professors Dodd Roberts, Alex Caughran, and George Prescott of the University of Maine for their support, advice, and encouragement throughout the conduct of the study and the preparation of this report as my doctoral thesis.

ABSTRACT

THE EFFECTS OF VISUAL TRACKING TRAINING UPON FIVE SELECTED ASPECTS OF THE SILENT READING PERFORMANCE OF FIRST GRADE PUPILS AS MEASURED BY EYE-MOVEMENT PHOTOGRAPHY

By Charles R. Thomas

This study was designed to test the proposition that visual tracking training, as represented by Visual Tracking, a published tracking program, would significantly improve first grade pupils' reading performance by improving perceptual processing efficiency. In addition, the study was designed to provide test-retest and inter-rater reliability data for first graders' eye-movement scores and to gather information regarding adaptive testing procedures to be used when photographing the eye movements of young children.

A posttest only factorial design was used to study the independent and interactive effects of two independent variables, visual tracking training and sex, upon the dependent variables, five eye-movement measures of reading performance.

The 115 first grade pupils were randomly assigned to the experimental and control groups and were regrouped for fifteen-minutes each day to participate in the experimental or control activity. Experimental group pupils used the Visual Tracking program and, at the same time, control group pupils participated in directed listening activities in separate rooms. The first grade teachers supervising the groups followed a weekly rotating schedule so that each teacher spent the same amount of time with each group.

At the conclusion of the twelve-week training period, eye-movements of all pupils were photographed with the EDL Reading Eye camera as they read silently a fifty-word test selection. To collect data for determination of test-retest reliability, retests were administered to forty-four pupils. The eye-movement films of all subjects were scored by the researcher and an assistant to gather data for determination of inter-rater reliability of the scoring.

Eye-movement scores of the ninety-six pupils remaining at the conclusion of the study were analyzed by factorial analysis of variance with the following results:

1. No significant differences were found between the experimental and control groups on any of the five eye-movement measures. The proposition that visual tracking training would significantly improve first

graders' silent reading performance by improving perceptual processing efficiency was not supported in this study.

2. The girls read with significantly fewer fixations and at a significantly faster rate with comprehension than did the boys. No significant differences were found on the sex variable for the other three eye-movement measures, namely, number of regressions, average span of recognition, and average duration of fixation.

3. No significant interaction effects were found between visual tracking training and sex.

Test-retest reliabilities ranging from .66 to .88 for the five eye-movement measures were similar to reliability coefficients reported in earlier studies. The test-retest reliability coefficients in this study were higher than most coefficients reported in previous studies using elementary school pupils.

Inter-rater reliability coefficients ranging from .82 to .98 led to the conclusion that scoring of eye-movement films has adequate reliability for use with groups of children.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

THE PROBLEM

Authorities in reading have emphasized the need for instruction and guidance during the pre-reading and initial reading stages to develop the basic skills in word perception that form the foundation for later progress in reading (Tinker, 1965, pp. 25-38). The technique of visual tracking was developed to aid teachers in that task.

Specifically, visual tracking training was designed to increase a learner's perceptual processing efficiency during reading. There is, however, little evidence regarding the validity of the technique. At first grade level where the training might logically have the greatest utility no evidence has been offered to indicate its value in improving reading performance.

I. BACKGROUND FOR THE STUDY

While the reading process has been defined in various ways, almost all definitions include some reference to the perceptual aspects of reading. Bergan (1965, p. 90) has emphasized the role of perception in reading as follows:

> It is axiomatic that reading is a perceptual task. The essence of the communication of ideas from the printed page is in the processing of information through the visual apparatus and the transformation of that information into meaningful thought units. This set of events is by definition perceptual.

Bergan (1965, pp. 90-91) also points out that perception is a generic term used to refer to many different situations and behaviors. Because perception is unobservable, he defines it as an intervening variable linked between the characteristics of the stimulus and the overt behavior of the organism.

Perception in reading, therefore, can be conceived of as a series of events occurring as the reader processes visual stimuli. Although the exact events occurring in perception are unknown, researchers are developing models of perception in reading (Geyer, 1968, 1969) and studying the influence of modality predilections--preference for visual, auditory, or tactile input channels--upon reading and other learning (Wepman, 1968; Bannatyne, 1968).

Visual Tracking, the published tracking program used in this study, is labeled "a self-instructional workbook for perceptual skills in reading" by the authors (Geake & Smith, 1962b); however, they give few insights into the nature of the perceptual skills affected by the program or the relationship of these skills to reading performance. This study was designed to investigate the effectiveness of the program for improving silent reading performance at first grade level.

When attempting to determine the effectiveness of any program for improving reading, an investigator must consider which measures of reading performance may serve most appropriately as the criterion or dependent variables. Previous studies of training with Visual Tracking (Geake & Smith, 1964; MacIver & Geake, 1965) used global tests of reading achievement that yielded, in one case, rate and comprehension scores and, in the other case, a composite oral reading grade level score.

A test that yields a rate of reading score may be appropriately sensitive to the specific kind of training involved in visual tracking, but it is unlikely that tests yielding grade level scores based upon comprehension or a composite based upon several reading behaviors are appropriate to serve as criterion measures of the effectiveness of specific perceptual training. Such broad instruments are likely to be insensitive to experimental treatment effects because they reflect the influence of many factors or variables that are not directly related to the training received.

Eye-movement photography is considered by the researcher to be a more appropriate criterion measure of reading performance for this study than other tests of reading achievement i the following reasons:

1. The authors of Visual Tracking (Geake & Smith, 1962b, p. 1) indicate that the program was developed in part as a result of a research finding that erratic eye movements were common just prior to errors in oral reading. They hypothesized that faster discrimination of letters and words after tracking training would result in fewer such erratic eye movements. Although this hypothesis cannot be directly tested in the present study because of equipment limitations, eye-movement photography with the Reading Eye should provide a more sensitive measure of improvements in the readers' perceptual processing efficiency that result from visual tracking training.

2. Studies (Ballantine, 1951; Buswell, 1922; Gilbert, 1953; E. A. Taylor, 1937; and S. E. Taylor, Frackenpohl, & Pettee, 1960) designed to determine age changes in eye-movement patterns from primary grades through college level have reported similar growth patterns for eye-movement measures. Although the studies differed greatly in size of sample, in the numerical bases for reporting scores, and in the particular cameras used, all indicated that the greatest amount of growth in eye-movement performance occurred at the earliest grade levels tested; that is, the growth curves for

each of the eye-movement measures rose steeply in the first three grades and then leveled off gradually at the upper elementary, secondary, and college levels.

In a recent comprehensive study (S. E. Taylor et al., 1960), eye-movement measures taken at the mid-year point of each grade level showed the greatest amount of change between the first and second grade testings. For fixations, the change between mid-year of first grade and mid-year of second grade was 37.4 per cent of the total change from grade one through college level. Figure 1 illustrates the percentage of change in the reduction of fixations for each grade level in the study. Figure 2 pictures the growth curves for each of the eye-movement measures. These two figures clearly indicate that the period between the middle of first grade and the middle of second grade is a time of rapid development for eye-movement patterns. By extrapolation, it might be assumed that growth in eye-movement measures is equally great during the first half of grade one.

First grade is a period of rapid growth in reading, and it is evident that mid-year measures of first graders' eye movements reflect this growth in reading performance. According to Tinker (1946, p. 113), ". . . Eye-movement habits are very flexible and appear to adjust themselves readily to any change in the perceptual processes involved in reading."

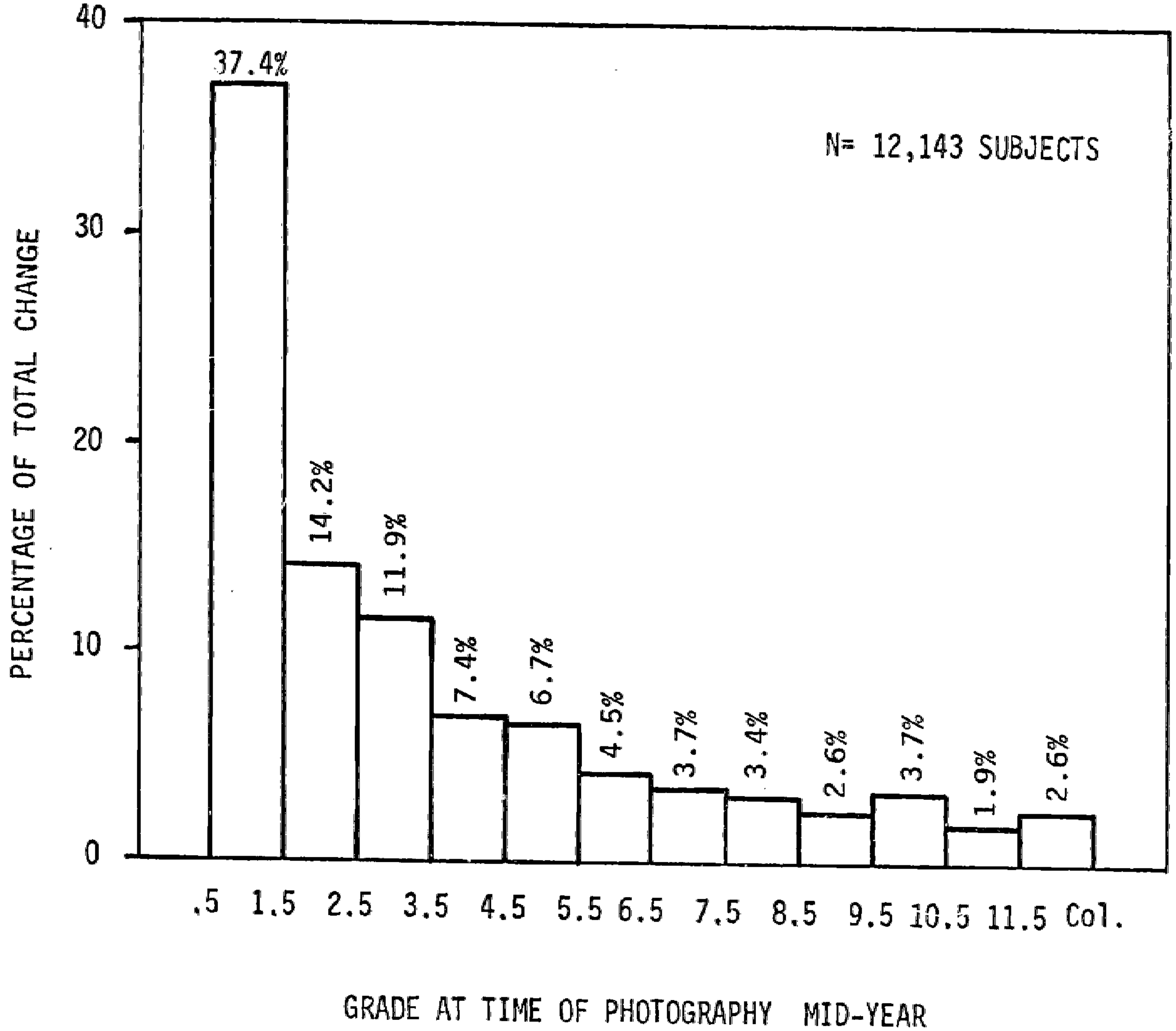


Figure 1. Pe-centage of total change in fixations that occurred between test ngs at each grade level with eye-movement photography. (Adapted from S. E. Taylor et al., 1960, p. 17.)

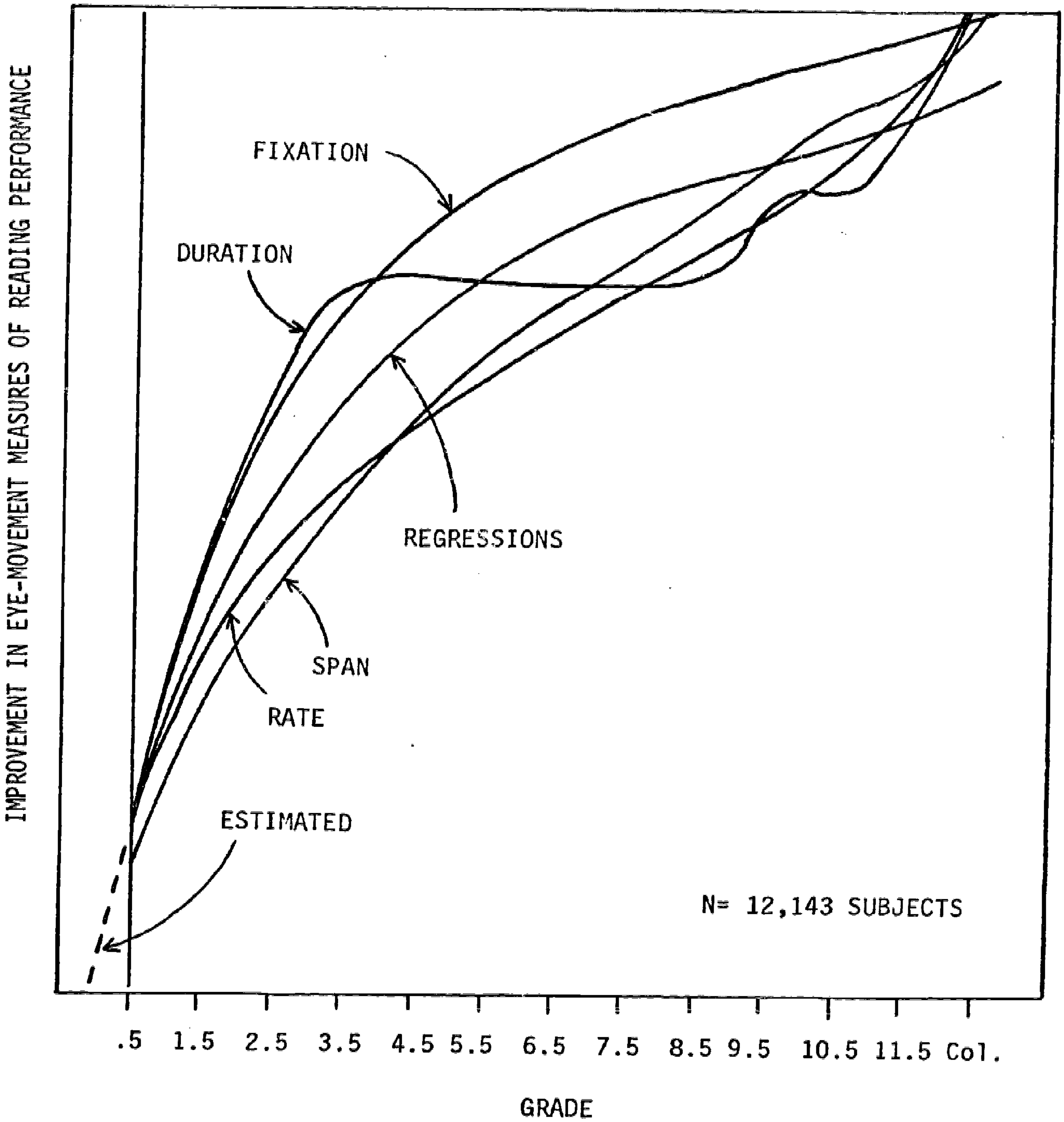


Figure 2. Curves for change in number of fixations, number of regressions, span of recognition, duration of fixation, and rate with comprehension determined by mid-year eye-movement photography of subjects from grade one through college level. (Adapted from S. E. Taylor et al., 1960, p. 12.)

It is likely, therefore, that eye-movement patterns of the subjects in this study would reflect changes in reading performance caused by the experimental or control treatments.

II. NEED FOR THE STUDY

The justification for this study lies in part in the implications it should have for the reading curriculum at first grade level. Visual tracking training appears to be a valuable addition to the procedures used in reading programs at primary grade levels. Because one cannot, however, depend solely upon intuitive judgments when making curricular decisions, there is a need for additional empirical evidence to support or refute the proposition that visual tracking training will significantly improve reading performance by improving a reader's perceptual processing efficiency. This study was designed to gather data that can provide such evidence.

The importance of this study also lies in the contribution it will make to the literature on eye-movement photography. Reviews by Tinker (1936a, 1946, 1958) indicate that more than two hundred studies of eye-movement photography have been conducted since 1901. The only investigations that involved first grade children were five studies that developed norms for eye-movement measures from first grade through college level. The information gathered in this researcher's study regarding reliability of the eye-movement measures, inter-rater reliability of the scoring, and adaptive procedures for the photography with first grade subjects is, therefore, needed to add to the existing literature on eye-movement photography.

III. STATEMENT OF THE PROBLEM

The purpose of this study was to determine, for a selected population of first grade children, the effectiveness of visual tracking training as represented by Visual Tracking, a published tracking program. To make this determination, the following problem was investigated: What are the effects of visual tracking training upon five selected components of first graders' silent reading performance measured by eye-movement photography? These five components are: (1) number of fixations, (2) number of regressions, (3) average span of recognition, (4) average duration of fixation, and (5) rate with comprehension.

The following hypotheses were tested in order to answer the primary question posed in this study:

1. There will be no significant differences between the mean performances of the experimental and control groups

on the following eye-movement measures:

a. number of fixations

b. number of regressions

c. average span of recognition

d. average duration of fixation

e. rate with comprehension

2. There will be no significant differences between the mean performances of subjects of either sex on any of the five eye-movement measures.

3. There will be no significant interaction between sex and visual tracking training on any of the five eye-movement measures.

This investigation was also designed to gather data regarding the test-retest and inter-rater reliabilities of the five eye-movement measures and to gather information regarding adaptive testing procedures needed to obtain the best results when photographing eye movements of first grade children.

IV. LIMITATIONS OF THE STUDY

Interpretations and generalizations of the findings of this investigation are subject to the following limitations:

1. The population investigated in this study was

limited to the entire first grade of one elementary school in Maine. Generalization of the findings from this study to first grade children in other schools is limited by the extent to which the first grade population is dissimilar to other populations.

2. This investigation was limited to changes in five components of silent reading performance that can be measured by eye-movement photography.

3. The tracking training used in this investigation was limited to *Visual Tracking*, a published program. Generalizations about performance with other visual tracking programs are limited to the extent that the materials are dissimilar to *Visual Tracking* in terms of content and procedure.

4. Conclusions based upon statistical treatments of the test data are limited by the extent to which the eye-movement measures and scoring have adequate reliability.

V. DEFINITIONS OF TERMS USED

Terms defined in this section are listed in alphabetical order to facilitate their use. Terms marked with an asterisk are illustrated in one or more of the appendices.

Average duration of fixation. The average duration of fixation is the mean time required for a reader's fixations. It is computed by dividing the total reading time for a selection by the total number of fixations required by the reader. For the Reading Eye, total reading time in seconds is determined by measuring the scored portion of the eye-movement film on a chart supplied with the camera.

Countable lines. Countable lines are the middle seven lines of print on a nine-line test selection card. In scoring Reading Eye films, the eye movements corresponding to reading the first and last lines of print are not counted in determining a subject's total score.

Easier test selections.* The easier test selections consist of four stories written by the investigator for use with the Reading Eye in this study. The stories were printed on cards and read by pupils who were unable to read the Reading Eye test selections with adequate comprehension.

Fixation.* A fixation is an eye-stop during reading. On an eye-movement film a fixation appears as a vertical line.

Inter-fixation movement.* An inter-fixation movement is a movement of the reader's eyes from one fixation point to another. On an eye-movement film an inter-fixation movement appears as a short horizontal or slanted line.

Perception time. Perception time represents the summation of all fixation durations during a reading. Perception time was reported in some of the earlier eye-movement studies that used cameras which permitted measurement of the precise duration of each fixation.

Rate with comprehension. Rate with comprehension refers to the rate of reading in words per minute for a selection on which a subject has evidenced adequate comprehension by scoring at least 70 per cent on a brief oral test of the story content.

Reading Eye.* The Reading Eye is a compact table model camera that photographs eye movements by the corneal reflection method. The Reading Eye was developed by Stanford E. Taylor of Educational Developmental Laboratories in 1959 to replace the bulky Ophthalmograph that had been in use since 1932.

Reading Eye test selections.* Reading Eye test selections are graded stories provided with the Reading Eye. A card containing one of these printed selections is inserted in a bracket on the camera to be read by the subject during the eye-movement testing.

Regression.* A regression is a fixation that follows a right to left inter-fixation movement of the eyes.

Return sweep.* A return sweep is a right to left movement of the eyes from the end of one line of print to the beginning of the next line. On an eye-movement film a return sweep can usually be distinguished from a right to left inter-fixation movement by the greater length of the return sweep line.

Span of recognition. A reader's span of recognition represents the number of words which he can process as a result of a single fixation. Since the span of recognition cannot be determined directly by inspection of an eye-movement film, an average span of recognition is computed by dividing the total number of words read by the total number of fixations made in the reading.

Visual tracking. In this study, a visual tracking task is one in which a learner visually scans lines of printed symbols and responds to individual symbols as directed. For example, he may circle or cross out the symbols, follow the symbols with a finger, or respond orally.

Visual Tracking program.* Visual Tracking is a published program in which the learner moves through a paragraph of nonsense words line by line and crosses out letters of the alphabet in sequential order. The program was designed to develop perceptual skills in reading.

VI. ORGANIZATION OF REMAINING CHAPTERS

Chapter II presents a review of the significant literature concerning visual tracking training, perceptual processing in reading, and the reliability and validity of eye-movement photography.

Chapter III describes in detail the design of the experiment and the procedures employed to conduct and evaluate the research. Also, procedures and findings of a related pilot study are summarized in Chapter III.

In Chapter IV, the data gathered in this study are analyzed. Finally, Chapter V presents a summary of the entire study, conclusions, implications, and recommendations for further research.

CHAPTER II

REVIEW OF THE LITERATURE

This chapter presents a review of the literature relating to the problem under investigation in this study. The review is organized around the following four general areas: (1) validity of visual tracking training, (2) perceptual processing in reading, (3) reliability of eye-movement scores, and (4) validity of eye-movement scores.

I. VALIDITY OF VISUAL TRACKING TRAINING

Studies Using the Visual Tracking Program

Geake and Smith (1964), authors of the _Visual Tracking_ program, have contributed two studies to the sparse evidence in the literature regarding the effectiveness of visual tracking training. The first report concerned a pre-publication pilot project with a group of forty retarded readers, aged eight to thirty. The authors reported that all but four subjects improved in perceptual speed and/or accuracy after completing the _Visual Tracking_ program. In discussing the same project in the _Visual Tracking_ manual, however, Geake and Smith (1962b, p. 2) stated: "Since no control group was available for comparison purposes, these results cannot be evaluated with exactness."

In addition to the lack of a control group, interpretation of the results of Geake and Smith's pilot project is hindered by two other serious problems. First, no statistical tests were used to determine the significance of the observed increases in test performance.

Second, the test used to measure perceptual speed and accuracy was inappropriate for some subjects in the pilot project. The Thurstone Perceptual Speed test (formerly known as Identical Forms) was designed for use in Grades 9-16 and in industrial employment (Buros, 1965, pp. 857-858) but was administered to subjects in the pilot project who were as young as eight years of age.

In summary, the absence of a control group, the assumption that observed gains in speed and accuracy represent real increases in subjects' skill without testing those gains for statistical significance, and the use of the Perceptual Speed test with some subjects younger than those for whom it was designed, all indicate that little confidence can be placed in the pilot project findings.

The second study (Geake & Smith, 1964) involved the use of Visual Tracking with children in second through fifth grades in a private school. At each grade level a class was divided into experimental and control groups on the basis of pretest reading scores. The Gates Advanced Primary Reading Test was used for second grade and the Gates Reading Survey was used for third, fourth, and fifth grades.

In each class the experimental group subjects completed the entire set of 160 Visual Tracking exercises over a three to four month period. During the practice sessions, the control group pupils in each class worked on unspecified "reading related" activities, presumably in the same rooms with the experimental groups. Both groups received regular classroom instruction in reading during the conduct of the program.

Results of posttesting in second grade indicated that the experimental group improved significantly more than the control group on word recognition but that differences between mean paragraph reading scores were not statistically significant. In the third and fifth grades, the experimental group showed significantly greater gain in mean scores on rate of reading and comprehension. No significant differences existed at fourth grade.

After one month, the pupils in third and fourth grades were retested with a third equivalent form of the Gates Reading Survey. At that time, both the third and fourth grade experimental groups showed significantly greater gains in rate and comprehension than did the control groups.

A follow-up test, presumably with another alternate form of the <u>Gates Reading Survey</u>, was administered to the third and fifth grade groups when they returned from summer vacation. By this time, some five months following the end of the tracking experiment, all significant differences between experimental and control group subjects in third and fifth grades had disappeared.

Despite the gains reported in favor of the experimental groups in the second study by Geake and Smith, interpretations of the findings are subject to these limitations.

First, because the study was inadequately reported, it is not possible to determine the extent to which extraneous variables were controlled or the type of statistical analysis employed.

Second, regarding the absence of a significant difference in mean scores for paragraph reading at second grade, Geake and Smith (1964, p. 63) wrote:

> This finding supported our expectation that the skills developed by the tracking exercises are immediately applicable to word recognition skill, and less directly to reading comprehension.

The above statement reflects a view similar to the position stated earlier, namely, that tests measuring reading comprehension may not be sensitive to the effects of the tracking training. Nevertheless, when the experimental groups did make significantly greater gains in comprehension at other grade levels, these findings, which were also contrary to the authors' expectations, were accepted without comment as true measures of treatment effects.

Third, the experimental and control groups at each grade level were called "equivalent" by the authors (Geake & Smith, 1964, p. 63), but the method by which this equivalence was achieved was not stated. Matched pairs could not have been used because there were unequal numbers of subjects in each group at each grade level. Although it was stated that subjects were assigned to groups on the basis of pretest scores, it is not clear how intact classes of twenty-three to twenty-five children could be so assigned to make the groups equivalent.

Fourth, the experimental group pretest means for rate of reading are from two to six words per minute higher than the control group means at each grade level. It is not clear that the initial mean differences, though slight, were tested for statistical significance. Another aspect of this problem is that the table which presents the data from this study does not include the pretest scores for reading comprehension at third, fourth, and fifth grades. The same table is presented in the Visual Tracking manual (Geake & Smith, 1962b, p. 3), but the comprehension pretest scores for third, fourth, and fifth grades are not listed in that source either. The reason for the omission is unknown.

Fifth, the specific instruction provided the control groups during the training session was not described. Since it is possible that both groups in each class worked on their respective materials at the same time in their regular classrooms, it is possible that contamination existed because of exposure of subjects in each group to the other group's materials. Also, it is possible that the performance of the experimental groups was affected because they received special treatment, i.e., they received a new material, Visual Tracking programs.

In a third study, MacIver and Geake (1965) used the Visual Tracking program with twenty-seven pupils in a public school fourth grade class. The Iowa Test of Basic Skills and Gray's Oral Reading Paragraphs were administered to the experimental group and to two other intact fourth grade classes that served as a control group. Pupils were given approximately five hours of tracking practice over an eight week period. Visual Tracking was used as a supplemental program so regular reading instruction was continued with the pupils.

At the end of eight weeks, the pupils were retested with the same tests. A mean gain of fourteen months in oral reading for the experimental group was statistically significant at the .01 level while a mean gain of six months for the control group was not statistically significant. On the vocabulary subtest of the Iowa Test of Basic Skills, both groups made gains statistically significant at the .05 level, five months gain for the experimental group and three months gain for the control group. On the comprehension subtest of the Iowa Test of Basic Skills no statistically significant gains were made.

The MacIver and Geake (1965) study appears to contain two weaknesses. First, the research design and procedure provided only limited control over relevant variables that could easily account for observed differences between the experimental and control groups.

For example, the senior author's intact class was selected to receive the experimental treatment. If it was not possible to assign individual students to the experimental or control groups at random, the determination of which intact class would receive the experimental treatment should at least have been randomly made (Campbell & Stanley, 1963, p. 217). Since it is difficult to control for effects caused by differences among the teachers when they teach only one class, this factor, an aspect of "intrasession history" (Campbell & Stanley, 1963, p. 184), will not be unknowingly manipulated in favor of the experimental group if the experimental treatment is assigned to one class at random.

Second, pretest-posttest differences in each group were tested for statistical significance. Instead, differences between the gain scores of the experimental and control groups should have been tested for significance if the t-test was used. An alternative would have been to use analysis of covariance with pretest scores as the covariate (Campbell & Stanley, 1963, p. 184).

Summary

Analysis of the three studies that used Visual Tracking suggests that the findings should be interpreted with caution because the studies were poorly reported and inadequately controlled. It is clear that the effectiveness of training with Visual Tracking as a means of improving reading performance needs to be researched further with studies that are more adequately designed.

II. PERCEPTUAL PROCESSING IN READING

Closely related to the question of the effects that visual tracking training has had upon reading performance in previous studies is the question of the relationships that exist between visual tracking behavior and perception in reading. In order to attempt to answer the latter question, it is necessary to explore more fully the nature of the perceptual process in reading.

<u>General Nature of Perception in Reading</u>

As the various aspects of perception are considered, it is important to keep in mind that, "Perception is a generic term which represents many different kinds of situations and behaviors (Bergan, 1965, p. 90)." As a result, the findings and conclusions of researchers vary widely. Conclusions regarding perception depend upon knowing whether the investigation concerned adults or children, good readers or poor readers, tachistoscopic presentation or sustained reading, or sight words or unfamiliar words.

Another difficulty encountered in discussing perception in reading is the fact that writers have differing views of the limits of perception and use the same terminology in differing ways. For example, in one text (Smith & Dechant, 1961, p. 23) reading is broadly defined as "interpretation of the printed page." It is further stated that "Interpretation requires both <u>recognition</u> and <u>perception</u>." Other writers (Anderson & Dearborn, 1952; Geyer, 1968, 1969;

Tinker, 1965), however, make no such distinction but include recognition as one aspect of the perceptual process that begins with sensory stimulation and ends with a response.

The difficulty with terminology is further illustrated by the following example. Dember (1964, pp. 16-26), in referring to studies of perception, listed four perceptual tasks: (1) detection, e.g., "Signal when you see something,"; (2) discrimination, e.g., "Is this light brighter than that one?"; (3) recognition, e.g., "Was the word you saw 'car,' 'cat,' or 'got?'"; and (4) identification, e.g., "What was the word you saw?". Each of the perceptual tasks from detection to identification requires that more information be obtained from the stimulus than the task at the level below it.

On the other hand, Tinker (1965, pp. 9-12), referring to perception in reading, reversed the order of the last two tasks. He described word identification, i.e., achieving correct pronunciation regardless of the amount of meaning present, as a lower level of perception than word recognition, i.e., accomplishing correct pronunciation and associating the sound with a previously known meaning of the word. Other writers use the terms interchangeably.

It is concluded that considerable differences exist regarding the definition and nature of the perceptual process in reading. According to Geyer (1969, pp. 7-9), the "classic view" of this process is based upon the span of attention experiments popular at the turn of the century. Reading, according to the "classic view," is seen as a series of tachistoscopic presentations flashed to the brain by the saccadic movements of the eyes: a unitary phenomenon in which all elements of the visual field are perceived instantaneously and simultaneously.

Word Perception in Children's Reading

One of the results of holding the "classic view" of perception in reading was described by Tinker (1965, p. 27):

> When the researches of Cattell and Erdmann and Dodge revealed that adults tend to read words as units rather than letter by letter, the word method received further impetus, for it was assumed, apparently wrongly, that young children perceive words in the same way adults do. For a time, teachers discarded all forms of word analysis. Each and every word was to be learned as a sight word by viewing it and repeating its sound many times. It was soon discovered, however, that use of the look-and-say method by itself is ineffective: it depends too much on guessing, which leads to many errors in word perception. Further, most writers now agree that only a few young children perceive and recognize words by the total word structure.

Yet, little is known about children's developmental progress in word perception in reading. The common assumption that young children perceive words in the same way adults do may have limited the research in that area. According to Vernon (1957, p. 22), "Much of the work on recognition in children's reading is speculative and inconclusive."

Apparently, children learning sight words by the look-and-say method recognize words by using a combination of word length and one or more familiar letters which represent the whole word. This recognition of words on the basis of the initial letter or other limited cues often results in confusions among words having similar length and the same cue letters in common.

Vernon (1962) agreed that the combination of general word shape or length and one or two letter characteristics can only serve to enable a learner to recognize at a glance his first words. She (Vernon, 1962, p. 108) added:

> But sooner or later, in order to perceive the essential structure of words he has to learn the characteristics of isolated letters and the manner in which they are combined in different words. This necessitates differentiating letters which are often similar in shape, particularly the reversed letters "b" and "d" and "p" and "q" and the inverted letters "n" and "u." He then has to learn that each of the letters has one or more associated sounds; and that the sound of the whole word is made up of the appropriate letter sounds in correct order.

Tinker (1965) reviewed the literature on word perception in adults and children and reported that children do not begin reading with the adult's proficient techniques of word perception. Instead, children develop perceptual skills slowly through experience and instruction. He (Tinker, 1965, p. 26) concluded:

> It would seem that children may have difficulty in learning to read until they are able and willing to perceive the small dissimilarities of form that differentiate letters and words from each other.

The need for a young reader to learn the characteristics of individual letters and to differentiate similar letters by perceiving minor differences in form may be partially fulfilled in the Visual Tracking program. According to the authors (Geake & Smith, 1964, pp. 62-63), the task in Visual Tracking involves the practice of a fundamental skill, discrimination, which involves memory, accuracy, and speed, and two other perceptual skills, tracking, i.e., ability to stay on one line of print, and orientation in space, i.e., left-to-right movement. Geake and Smith (1964, p. 62) describe the discrimination process as follows:

> As the student races through the paragraphs of nonsense words in search of the letters, he is practicing the perceptual skills. In order to finish successfully and in good time he must stay on the correct line of print, proceed from left to right and discriminate the individual letters (thus using memory and accuracy) at the highest speed at which discrimination is possible for him at his present level of skill.

The authors of Visual Tracking also point out that the required discrimination is complicated by a controlled number of competing stimuli. For example, a _b_ will have a _d_ near it on the same or adjacent line. By gradually reducing the degree of letter and line spacing in the later exercises, the discrimination task is made increasingly difficult.

In view of the fact that letter reversals and confusions among letters of similar shapes are commonly reported in studies of normal readers until ages seven or eight and longer in backward readers (Vernon, 1957, 1959, 1966), the letter discrimination practice provided by Visual Tracking would appear to have value for many first grade pupils.

Eye Movements and Perception

According to the "classic view" (Geyer, 1969, pp. 7-9), perception in reading is a unitary phenomenon in which all elements of the visual field are perceived instantaneously and simultaneously. With such a view of perception, it is difficult to conceive of any role for eye movements beyond their obvious progression along a line to allow the "photographing" of segments of print.

Since it has been demonstrated that no clear retinal image is produced when the eye is in motion, it has been concluded by some that no perception occurs during inter-fixation movements. For example, Tinker (1965, p. 15) wrote: "Since research has shown how the eyes move in reading and that perception occurs only during the pauses, . . ." Also, S. E. Taylor (1960, p. 36) stated: "Fixation or fixation pause refers to that period in reading a line of print during which the eyeball is held stationary for a short time and during which perception takes place."

<u>Perception as a continuous process</u>. In contrast to the "classic view" of perception in reading is the view expressed by Woodworth and Schlosberg (1954, pp. 506-507):

> As a matter of fact, the perceptual processes undoubtedly go on during the saccadic movements, even though retinal stimulation is not effective with the moving eye. Reading is a continuous process in that the perceptual development of meaning goes on steadily. Perhaps one can think of it as a continuous production process, a machine into which the raw material is tossed by the shovelful. The output will be continuous, as long as there is some raw material in the works. This analogy has another similarity to reading; the rate of input will usually be limited by the rate at which the machine processes the raw material and not vice versa. In a similar fashion, the eye movements adjust to the rate at which O [the observer] is digesting the sensory input.

Bergan (1965, p. 96) has similarly written: ". . . perception, as a series of events, involves acts which occur after information is received by the receptors as well as events occurring during the reception process itself." In addition, Geyer (1969, p. 18) wrote:

> There is wide acceptance today that perception is not a unitary process and that the sensory and response phases are more independent of one another than had long been assumed. Recognition of stimulus elements continues after the cessation of the physical stimulus.

The most significant evidence for the concept of perception in reading as a series of events that occur over time is the memory after-image effect frequently reported in tachistoscopic studies of letter and word identification (Geyer, 1969; Neisser, 1967; Vernon, 1962). One aspect of this effect is that subjects can recognize longer phrases from a very brief tachistoscopic exposure than they apparently can during a fixation in reading. Another aspect is that subjects in such experiments consistently report that they saw all the letters in a briefly presented random array, but that they forgot some before they could be reported.

To control the amount of time a subject has to process the stimulus from his memory after-image, Gilbert (1959) interrupted the after-image through projection of nonsense material at varying intervals following tachistoscopic flash. He found that shortening the period left free from interfering stimuli had the same effect upon the subjects' responses as did shortening the duration of tachistoscopic exposure with no interruption.

As a result of his experiment, Gilbert suggested that individual differences in speed of processing visual material is an influencing factor in both the span of perception and the duration of the fixation pauses in reading. In conclusion, Gilbert (1959, pp. 13-14) stated:

> These data suggest the possibility that some readers may use part of their fixation time to avoid interference from a new stimulus during the period they need free for processing the visual stimulus. In other words, part of the fixation time may be preventative in nature.

The research of Geyer (1968) and the model of perceptual processing in reading he developed (Geyer, 1968, 1969) can be combined with Gilbert's findings to provide a plausible explanation for the role of eye movements in reading. Viewing perception in reading as a series of events that take time, it is possible to view the role of eye movements as a means of keeping the interacting perceptual systems working smoothly. That is, eye movements may serve to make adjustments that maintain balance when one system gets ahead of the other.

On the basis of his research, Gilbert (1959) suggested that part of a reader's fixation time may be preventative in nature. That is, part of the fixation time is used to avoid interference from a new stimulus during the period the reader needs free for processing the prior stimulus. Similarly, Geyer (1968), found that some fixations of subjects in his study of the temporal eye-voice span in oral reading appeared to be for the purpose of marking time only. In summary, Geyer (1968, p. 52) stated:

> These graphs showed clearly that in reading under the conditions of this experiment, a significant portion of the eye-movement pattern is related to a necessity to balance temporally input and output systems and that at points of error or at voice pauses, the eyes must take some corrective action in order to maintain or re-establish the temporal balance.

If eye movements in reading serve primarily to balance the systems operating in the perceptual process, then additional support is provided for the commonly held view (Spache, 1960; Tinker, 1946, 1958, 1965; Woodworth & Schlosberg, 1954) that eye movements are limited by central processes of perception (recognition, assimilation, comprehension, etc.) and merely reflect the quality of the reading performance. The movements of the reader's eyes do not cause the reading to be higher or lower in quality.

<u>Role of oculomotor behavior</u>. On the other hand, E. A. Taylor (1959) and S. E. Taylor (1960, 1965) reason that eye movements are not completely at the bidding of the

central processes. On the basis of their extensive work with eye-movement photography, visual functioning, and instruments for improving reading, both authors conceive of an interdependent relationship between eye movements and the central processes. In support of the interdependence view, S. E. Taylor (1965) indicated that research has shown eye-movement or oculomotor patterns to be affected less than would be expected by changes in difficulty of material, reader's purposes, variations in subject matter, and the physical condition of the reader in terms of fatigue. He argued that the reader's oculomotor pattern is a habitually acquired activity that does affect the quality of the reading performance.

In describing the oculomotor development of young children, S. E. Taylor (1965, pp. 194-196) reported:

> During the readiness stage of reading instruction, he [a pupil] is given a minimal amount of assistance in visual discrimination and usually only verbal directions to help him form his directional attack. Throughout the beginning grades, the struggle with word identification and recognition encourages excessive and random ocular movements. In addition, the child may be handicapped by inadequate visual acuity or accommodation, poor binocular coordination, or lack of ocular facility or motility.
>
> Thus, by constant trial and error, a beginning reader unconsciously evolves a habitual oculo-motor activity that enables him to realize a reasonable degree of meaning from print, but one that may have

an inhibiting effect upon the development of reading efficiency in the later grades.

If Taylor's view is valid, then it is appropriate that some means of training oculomotor behavior be used in the early stages of reading instruction. The motor practice provided by Visual Tracking may have value for improving oculomotor patterns because it involves tracking, i. e., the ability to stay on one line of print, left-to-right progression, and precise control of the eyes as letters are rapidly discriminated.

Considerable disagreement exists, however, regarding the relationship between oculomotor ability and reading performance. To determine the relationship, Gilbert (1953) studied the oculomotor ability, as measured by eye-movement scores obtained from fixating upon lines of unevenly spaced digits, of 473 pupils in first through ninth grades. He also correlated eye-movement scores for fixating digits with eye-movement scores for prose reading in second through ninth grades. These correlations for number of fixations, number of regressions, and average duration of fixation ranged from .31 to .71. Most correlations, however, were in the .50's and .60's. For forty-two college students also studied, correlations between eye-movement scores for fixating digits and for reading prose ranged from .54 to .63.

Gilbert's overall conclusion was that with respect to fixations and regressions, there is a substantial relationship between oculomotor control involved in the simple motor activity (fixating digits) and oculomotor control involved in reading prose. For duration of fixation, the relationship is not so pronounced.

On the basis of the wide individual differences found in oculomotor control at all levels, Gilbert also concluded that ability to control the eye is not the pure product of experience in reading. He argued that since some children come to first grade capable of better oculomotor control than some others can demonstrate after seven or eight years of reading experience, eye-movement behavior does not necessarily stem from habits acquired in learning to read (Gilbert, 1953, p. 214).

Tinker (1958, 1965), in reviewing Gilbert's study of oculomotor control, criticized the findings on several counts. His major criticism was that Gilbert's measure of simple motor activity, namely, fixation of digits, was not really a measure of motor coordination but of simple reading.

He pointed out that since numerals are symbols for words and differ only in space required for printing, the digits were actually read by Gilbert's subjects just as words would be read. After noting that all subjects in Gilbert's study read an identical prose selection, Tinker (1965, p. 78) concluded:

> Eye movements in reading such material should be little influenced by comprehension factors. So it is not surprising that oculomotor patterns (fixation and regression frequency, pause duration) for reading the digit series and the easy prose turned out to be somewhat similar. Undoubtedly they are to some degree similar reading situations.

In an investigation of college students' oculomotor behavior in reading, Tinker (1938) compared motor coordination or accuracy of fixation, as indicated by speed of convergence-divergence movements and by the number and extent of ocular adjustments, with speed of reading scores on a standardized test and with eye-movement scores based upon prose reading. He found no significant correlation between any of the measures of oculomotor coordination and the measures of reading proficiency. He also found that only when extremes of the group were compared did a slight and consistent relationship between motor efficiency of the eyes and reading performance appear.

Tinker (1936a, 1946, 1958) also reviewed many studies in which an attempt was made to improve reading performance by training eye movements and failed to find that such training is either necessary or desirable. He (Tinker, 1965, p. 109) concluded:

> Many so-called procedures for training eye movements or for controlled reading result in improved reading efficiency, either in speed or in both speed and comprehension. This improved reading status is reflected in modified oculomotor patterns. But the improvement obtained by eye-movement training, with or without elaborate apparatus, is not greater than that resulting from well-motivated reading alone. Furthermore, there is no adequate evidence that training eye movements as such improves reading. Examination of experiments concerned with pacing eye movements and controlled reading reveals that they usually involve other training techniques as well and are never divorced from increased motivation.

It must be kept in mind, however, that most attempts to improve oculomotor control or eye-movement patterns have used high school and college age subjects. Little has been done at the beginning reading stage where, according to Gilbert (1953, pp. 228-230), inequalities in motor development are likely to assert themselves. Although immature oculomotor coordination may play a relatively minor role as a cause of poor reading or as a deterrent to good reading, Gilbert indicated that the subject of motor efficiency and its improvement is worthy of further study.

In regard to improvement of the oculomotor coordination of beginning readers, Gilbert (1953, p. 230) wrote:

> There is the possibility that simple training exercises can be introduced in the first grade to go along pari passu with other reading instruction without detriment to the recognition-comprehension processes.

The effectiveness of one training material, Visual Tracking, in improving the silent reading performance of first grade children will be investigated in this study.

Cue Reduction

Before closing this section, another relationship between eye movements and perception in reading must be explored. According to Anderson and Dearborn (1952, p. 162):

> Cue-reduction refers to the refinement of a skill and to the elimination of waste motion. In the case of reading, the concept of cue-reduction applies not only to the motor side of the performance but also to the sensory side, in that, with time, fewer cues from the printed page or less of the original pattern of stimulation is required to get the meaning. The efficiency of the eye movements which characterizes skillful silent reading is a function of cue-reduction.

Whereas it was proposed by Geyer (1969) that eye movements may reflect the extent to which balance exists among the operating perceptual systems, Anderson and Dearborn have proposed that eye movements reflect the refinement of perceptual skill through cue reduction. If, however, perception in reading represents the operation of a series of events over time rather than an instantaneous and unitary process, how is cue reduction accomplished and how might visual tracking training facilitate the process? The answer may lie in a portion of Geyer's (1968, 1969) model of perceptual processing.

The first phase in Geyer's model consists of two sensory systems, an initial sensory input system and a sensory organizational system. The purpose of the sensory input system is to scan the visual image within the retinal area at a fixed rate. Then, immediately upon input from the scanning system, organizational and associational processes begin. It is at this point that the development of meaning begins.

The purpose of the sensory organizational system is to organize and transform the visual image into higher order units prior to transfer to iconic storage, a short-term memory system. Although Geyer (1969, p. 32) indicated that the transformation is presumably phonemic, he did not describe the process by which the transformation takes place.

For more details, we must turn to the work of Gibson (1963) who pointed out that though the whole area of information processing has never been studied developmentally, the formation of larger units may be a very important aspect of perceptual development. She (Gibson, 1963, p. 190) added:

> Here would seem to be one more type of perceptual learning, the formation or registration of higher order units. It was suggested earlier that such units may be generated by "rules" which are progressively induced as samples of written words and their auditory correlates are encountered.

In a recent experiment (Gibson, Pick, Osser, & Hammond, 1962), it was found that nonsense letter groups with high phoneme-grapheme correspondence were reproduced more accurately than equivalent letter groups with low phoneme-grapheme correspondence at all durations of tachistoscopic exposure. The investigators reasoned that the difference could not have been caused merely by the familiarity of the letters because each pair of words used the same letters and letter clusters (ch, bl, etc.). Instead, they concluded that the differences must be attributed to higher-order graphic units, namely, the letter combinations of English writing that function as relatively stable units in phoneme-grapheme correspondences.

Gibson et al. (1962, p. 564) further wrote:

> While reading is based upon discrimination and identification of visual forms such as letters, it becomes, in the skilled reader, a process of perceiving "super forms," and these tend to be constituted (organized) by their relation to auditory-vocal temporal patterns.

In questioning how children progress from the differentiation of letters to the perception of words, phrases, and sentences, Gibson (1963) suggested that an inductive kind of perceptual learning occurs. Though not verbalized, this learning allows children to internalize the rules of English phoneme-grapheme correspondences. As these internalized rules are developed, they speed up perception by enabling the reader to organize the visual image in higher-order chunks.

The process of transformation into higher-order units is described in more detail by Miller (1956, p. 93):

> Since the memory span is a fixed number of chunks, we can increase the number of bits of information that it contains simply by building larger and larger chunks, each chunk containing more information than before. . . .
> In the jargon of communication theory, this process would be called recoding. The input is given in a code that contains many chunks with few bits per chunk. The operator recodes the input into another code that contains fewer chunks with more bits per chunk. There are many ways to do this recoding, but probably the simplest is to group the input events, apply a new name to the group, and then remember the new name rather than the original input events.

Geyer (1969) indicated that in reading unconnected letters or digits, relatively little recoding would be done by the sensory organizational system before transfer to iconic storage. In the case of prose reading, however, the letters would be transformed to higher order units prior to storage. He also suggested that initial processing at the sensory organization level involves recoding of letters to phonemes, Gibson's grapheme-phoneme units, or in advanced stages of reading, whole words.

It is concluded, therefore, that the effectiveness of Visual Tracking as a means of improving perceptual processing in reading may depend upon the extent to which cue reduction is a function of familiarity, as suggested by Anderson and Dearborn (1952), or some form of graphemic-phonemic recoding. The increased familiarity resulting from rapid discrimination of letters in Visual Tracking may facilitate the reduction of cues necessary for word recognition. On the other hand, if cue reduction is a function of recoding into higher-order units, Visual Tracking may facilitate perceptual processing only for those first grade pupils who are prevented from developing such units because of a more basic problem, namely, confusion among letter forms.

III. RELIABILITY OF EYE-MOVEMENT SCORES

Reliability Studies

Most of the research on the reliability of eye-movement scores was conducted in the 1930's using college students or pupils from the upper elementary grades as subjects and using cameras that are now obsolete. The earliest studies (Eurich, 1933a, 1933b; Frandsen, 1934; Litterer, 1932; and Tinker, 1936b) used the University of Minnesota eye-movement camera, an enormous piece of equipment permanently mounted on a ten-foot-long table at which the subject was seated. Two other studies (Broom, 1940; Imus, Rothney, & Bear, 1938) used the Ophthalmograph, a portable floor model camera. Only one study (S. E. Taylor, Frackenpohl, & Pettee, 1959) has reported reliability data based upon use of the Reading Eye.

All three cameras used in the reliability studies record eye movements by the corneal reflection method. Although the most obvious improvement in the cameras is the great reduction in size, the most important advantage of the Reading Eye is that the lights shining on the subject's eye are dimmed during the filming. With the lights dimmed, the reader experiences less distraction and discomfort.

By reference to Table I, where each of the studies reporting reliability data is summarized, it may be seen that considerable variation exists among the studies in terms of grade level of subjects, length and difficulty of the reading material, method of determining reliability, and in the reliability coefficients obtained for each eye-movement measure. Because of these variations in cameras, subjects, materials, and procedures, it is difficult to compare the studies directly. Rather, only general conclusions can be drawn.

The studies that determined internal consistency reliability through the split-half or odd-even methods (Frandsen, 1934; Litterer, 1932) obtained relatively high reliability coefficients. Except for two lower correlations of .77 and .78, internal consistency reliability coefficients ranged between .87 to .92, comparing favorably with the reliability of most standardized reading tests.

In general, test-retest reliability coefficients tend to be highest for the longest selections read. For subjects who read one hundred word selections, S. E. Taylor et al. (1959) reported reliability coefficients ranging

TABLE I

SUMMARY OF RELIABILITY COEFF[...]
FOR EYE-MOVEMENT SCORE[...]

Study	Number of Subjects & Level	Material Read	Method of Determining Reliability	Fixatio[...]
University of Minnesota Camera				
Litterer (1932)	71-College	14 lines-easy prose	split-half	.85(.92
	76-College	14 lines-harder prose	split-half	.63(.77
Eurich (1933a)	173-College	2 paragraphs - 61 & 63 words each	test-retest[b]	.74
Eurich (1933b)	100 4th & 5th gr.	3 paragraphs - 63,52, & 66 words each	test-retest[b]	.70-.87
Frandsen (1934)	66-College	1 para. scientific prose (length not stated)	odd-even	.77(.87
Tinker (1936b)	77-College	6 para. easy prose & 7 para. hard prose	test-retest[c]	.63-.84
		23 lines easy prose (Combination of 3 sel. each day)		.88
		38 lines hard prose (Combination of 4 sel. first day correlated with 3 sel. second day)		.78
Ophthalmograph				
Imus et al. (1938)	150-College	3 Ophthalmograph cards 50 words each	test-retest[b]	.61-.72
Broom (1940)	192-4th-6th grades	2 Ophthalmograph cards 50 words each	test-retest[c]	.79
Reading Eye				
S.E. Taylor et al. (1959)	30-College	5 Reading Eye cards 100 words each	test-retest[b]	.83-.91

[a]Reliability for total selection estimated by Spearman-Brown prophecy formula
[b]Same sitting
[c]Different sittings

TABLE I

SUMMARY OF RELIABILITY COEFFICIENTS
FOR EYE-MOVEMENT SCORES

l Read	Method of Determining Reliability	Reliability Coefficients			
		Fixations	Regressions	Fixation Duration	Perception Time or Rate
s-easy prose	split-half	.85(.92)[a]			.83(.90)[a]
s-harder prose	split-half	.63(.77)[a]			.83(.91)[a]
raphs - 61 & rds each	test-retest[b]	.74	.62	.83	
aphs - 63,52, ords each	test-retest[b]	.70-.87	.68-.87	.62-.69	
scientific (length not)	odd-even	.77(.87)[a]	.64(.78)[a]		.79(.88)[a]
easy prose & hard prose	test-retest[c]	.63-.84	.58-.83	.66-.89	.52-.83
easy prose tion of 3 sel.)		.88	.80	.82	.82
hard prose tion of 4 sel. correlated l. second day)		.78	.86	.85	.72
mograph cards s each	test-retest[b]	.61-.72	.59-.60		.59-.62
mograph cards s each	test-retest[c]	.79	.76		.78
Eye cards ds each	test-retest[b]	.83-.91	.84-.93	.86-.93	.90-.96

y Spearman-Brown prophecy formula

from. 83 to .96. For subjects who read combinations of selections totaling twenty-three and thirty-eight lines, Tinker (1936b) reported reliability coefficients ranging from .72 to .88. The reliabilities obtained by these authors were among the highest reported in any study.

On the other hand, reliability coefficients ranging from .59 to .72 for fifty-word selections in the study by Imus et al. (1938) were among the lowest reported in any study. Tinker (1936b, p. 745) in reviewing his findings concluded:

> The eye-movement measures, even for reading selections of five or more lines, have adequate reliability where group measures are concerned. This is true even for the first selection read before the camera.

Broom (1940, p. 208) expressed a similar view concerning the fifty word Ophthalmograph test selections:

> The reliability of the Ophthalmograph card tests for fixations, regressions, and speed is fairly satisfactory for group measurement, but it is such that those tests should rarely be used for individuals. . . .

On the other hand, Anderson and Dearborn (1952, p. 109) expressed a more conservative view after reviewing the research:

> When an adequate amount of material is used for the test before the camera, the various eye-movement

> measures yield reliability coefficients which compare favorably with most pencil-and-paper tests of reading achievement. A minimum of twenty lines of materials is recommended for eye-movement measurements when the object is to assess individual performance. Half this number of lines is sufficient for group measurements.

Morse (1951) compared the reliability of eye-movement scores in the literature with the reliability of ten standardized reading tests. After pointing out that the eye-movement reliability coefficients compared favorably with several of the standardized reading test reliability coefficients, Morse (1951, pp. 32-33) concluded:

> The remainder of the correlations [about 8 out of 25 coefficients listed] . . . are higher than the coefficients for eye-movement scores, but the typical paper-and-pencil test also contains considerably more material than the typical test before the eye-movement camera. The difference frequently amounts to a thousand words or more. Considering this difference, it is remarkable that the reliability coefficients for the different eye-movement measures turn out as high as they do.

Inter-Rater Reliability

Regarding the objectivity of eye-movement photography, S. E. Taylor *et al*. (1960, p. 2) wrote:

> Lastly, analysis of eye-movement photographs reveals data that is objective, with factors that are directly countable and measurable. Further objectivity results from the fact that the test records directly the subject's activity rather than a written or oral expression of this activity.

Apparently, other researchers have also viewed eye-movement photography as a completely objective technique because none has previously investigated inter-rater reliability of the scoring.

The experience of this researcher has indicated, however, that the procedures for scoring eye-movement films do include opportunities for scorer variation. It is concluded, therefore, that data on inter-rater reliability of the scoring is needed to determine the extent of such variations.

Summary

The available data indicate that reliability of eye-movement photography scores is generally considered too low for accurate measurement of individual performance, but that reliability is adequate for study of groups if the selections to be read are of comparable difficulty and of sufficient length. Unfortunately, investigators disagree regarding the number of words or lines of print that constitute sufficient length.

Apparently the one hundred-word, twelve-line (ten countable lines) Reading Eye test selections read by college students in the study by S. E. Taylor *et al.* (1959)

were of sufficient length because the reliability coefficients obtained were high. But, are the fifty-word, nine-line (seven countable lines) Reading Eye test selections used at first grade level long enough for the eye-movement scores obtained to be reliable? This question cannot be answered from the literature with any degree of assurance because no previous study has reported reliability coefficients for the eye-movement scores of first grade subjects. This study was, therefore, designed to provide such reliability data.

IV. VALIDITY OF EYE-MOVEMENT SCORES

Typicality of the Performance

In assessing the validity of eye-movement photography measures, it is common to question whether the subject's performance before the camera is a typical reading performance. To investigate this question, Tinker (1936b) had 134 college students read different forms of the Chapman-Cook Speed of Reading Test before the University of Minnesota eye-movement camera and away from it. He reported a correlation of .94 for reading rate and concluded that performance was the same with and without the camera.

In another study (Gilbert & Gilbert, 1942), forty-seven fifth grade pupils read different comparable selections before and away from the University of Chicago eye-movement camera. Although it was found that the pupils read slightly faster before the camera, differences in rate and comprehension were not statistically significant.

In summary, S. E. Taylor et al. (1959, p. 2) concluded:

> There seems to be little need to validate eye-movement photography, for the work by Gilbert and Gilbert, and Tinker conclusively established the fact that students read similarly before and away from an eye-movement camera.

Several investigators (Gilbert & Gilbert, 1942; Morse, 1951; Seibert, 1943; Tinker, 1936b) have emphasized, however, that it is essential to establish good subject rapport and to provide an adequate orientation to the eye-movement camera if a typical reading performance is to be obtained before the camera.

Validity Coefficients

The method which has commonly been employed to study the validity of eye-movement scores has been to correlate these scores with results on standardized tests of reading achievement. Since better reading performance is characterized by fewer fixations and regressions and by shorter duration of fixation, validity correlation coefficients between these measures and reading test performance are expected to be negative. On the other hand, since a faster rate characterizes better reading performance, validity coefficients for rate of reading are expected to be positive.

Early studies (Anderson, 1937; Eurich, 1933a, 1933b; Imus et al., 1938; Litterer, 1932; Tinker, 1936b) reported validity coefficients between several different paper-and-pencil reading tests and eye-movement scores ranging from -.02 to -.71 for number of fixations, from -.23 to -.41 for number of regressions, from -.05 to -.34 for duration of fixation, and from .32 to .71 for perception time or rate of reading.[1]

Tinker (1936b) explained the wide variation in validity coefficients in terms of the fact that the material

---

[1]All validity coefficients presented in this section are uncorrected. Many of the studies reported coefficients corrected for attenuation, but it is doubtful that the corrected correlations adequately represent reality. When the attenuated r's are corrected for the reliability of both the criterion test and the eye-movement measures, as S. E. Taylor et al. (1959, p. 6) did, there is a particular danger that the corrected validity coefficients will be inflated.

read before the camera is not the same as the criterion test in most cases. Also, he pointed out that since there are many specific reading skills rather than a general reading ability, expected intercorrelations between various reading tests range from .00 to .60.

Tinker (1936b, 1965) and others (Anderson & Dearborn, 1952; Morse, 1951) have indicated that validity coefficients for number of fixations and rate of reading are relatively high when the material read before the camera is strictly comparable to the criterion test. Generally, the comparable material is an alternate form of the same reading test. For example, in one study (Tinker, 1936b) college students read two selections totaling twenty-eight lines from the Chapman-Cook Speed of Reading Test before the camera. Using scores on the entire test read away from the camera as the criterion, validity coeffi- ranged from -.56 to -.71 for fixations, from .64 to .71 for rate of reading, and from -.08 to -.24 for duration of fixation.

On the other hand, alternate forms of the same test do not have to be used to obtain high validity coefficients for rate of reading. In a recent study (S. E. Taylor et al., 1959), ninety students, thirty each in the fourth, seventh, and tenth grades, were tested with the rate portion of the Diagnostic Reading Tests. Eye movements were photographed the following day as subjects read the regular graded Reading Eye test selections. Pearson r correlation coefficients between the rate portion of the Diagnostic Reading Tests and rate of reading as measured by eye-movement photography were .83 at fourth and seventh grades and .91 at tenth grade. According to the authors, these correlations support the conclusion that eye-movement photography can be used as a valid measure of rate.

In summarizing the validity data for eye-movement photography, Morse (1951, p. 37) wrote:

> High correlations require the use of comparable materials in the camera situation and in the criterion. When determined on this basis, the validity coefficients for fixation frequency and total perception time turn out extremely high. Fair validity will be found for regression frequency and low validity for pause duration.

Other Validity Evidence

In order to determine the value of eye-movement photography as a measure of reading ability, Anderson (1937) photographed the eye-movement of 174 college freshmen as they read materials varying in level of difficulty. His findings indicated that each measure of eye movements distinguished good from poor readers at each level of difficulty of material. Although the eye movements of both the good and the poor readers were influenced similarly by changes in the difficulty of the material read, the good readers were more flexible and able to modify their eye movements over a wider range than the poor readers. Anderson (1937, p. 30) concluded that this flexibility of eye movements following changes in central processes of apprehension and comprehension emphasizes the dependence of eye-movement behavior upon reading ability.

Walker (1933) photographed the eye movements of college freshmen who had been classified as good readers on the basis of high performance on two reading tests. Changes in eye-movement measures as the material read increased in difficulty led the investigator (Walker, 1933, p. 109) to conclude that comprehension is a very significant determiner of eye movements. That is, the eye movements of good readers are symptomatic of the ease or difficulty with which the reader comprehends the material.

Finally, on the basis of reviews of well over two hundred eye-movement studies, Tinker (1965, p. 111) concluded:

Eye-movement patterns are very flexible and apparently adjust themselves readily to any changes in the perceptual and assimilative processes involved in reading. It appears that eye movements merely reflect, or are symptoms of, efficient or poor reading performance.

Summary

Available data indicate that students read similarly before and away from an eye-movement camera when good subject rapport has been established and when an adequate orientation to the camera has been provided. In general, researchers reported that eye-movement measures can distinguish good from poor readers and adequately reflect the quality of a student's reading performance.

CHAPTER III

RESEARCH PROCEDURE

This chapter describes the design of the experiment and the procedures employed to conduct and evaluate the research. Also described in this chapter are the findings of a pilot study in which testing procedures were developed and four easier test selections written by the experimenter were studied.

I. RESEARCH DESIGN

This study utilized a posttest-only factorial design with two independent variables, treatment and sex. Comparisons were made on the treatment variable between mean scores of experimental and control group subjects; on the sex variable, between mean scores of male and female subjects.

The dependent variable was silent reading performance as represented by each of the following eye-movement measures: (1) number of fixations, (2) number of regressions, (3) average span of recognition, (4) average duration of fixation, and (5) rate with comprehension.

With the use of a factorial design, it was possible to study the independent and interactive effects of the two independent variables upon the dependent variable (Kerlinger, 1964, p. 325). A posttest-only design was used to avoid the following problems that might have resulted from the administration of an eye-movement photography pretest:

1. In the preliminary stages of this study, the first grade teachers estimated that 30 per cent of their pupils lacked the reading skill necessary to read the Reading Eye test selections with adequate comprehension. Thus, about 30 per cent of the pupils would have been eliminated from the study because they lacked a pretest score. The participating 70 per cent would have been biased in favor of the better readers while the remainder, for whom the tracking training might have had value, would have been unable to participate in the study.

2. It was possible that an eye-movement photography pretest would have a sensitizing effect upon the experimental group pupils and affect their posttest performance (Campbell & Stanley, 1963, pp. 188-191). The possibility of pretest sensitization existed because eye-movement photography is an unusual school activity

and because the testing would have been conducted by the investigator, an outsider. These two factors could have made the pupils or their parents partially aware of the experiment and its purpose. If so, then the experimental group pupils, whose training was more closely related to the performance required by the test, might have been affected more by the training than they were without a pretest.

The posttest-only design with complete randomization used in this study not only avoids problems associated with the pretest but also gives assurance that abilities on any variable are equally distributed between the experimental and control groups. Campbell and Stanley (1963, p. 195) state:

> For psychological reasons, it is difficult to give up "knowing for sure" that the experimental and control groups were "equal" before the differential experimental treatment. Nonetheless, the most adequate all-purpose assurance of lack of initial biases between groups is randomization. Within the limits of confidence stated by the tests of significance, randomization can suffice without the pretest.

An illustration of equalizing the experimental and control groups by random assignment is provided by a comparison of mean chronological ages for the two groups. The mean chronological age of the fifty-eight experimental group pupils was 85.51 months as of March 1, 1969. The mean chronological age of the fifty-seven control group pupils at the same time was 85.07 months. The difference of .44 months between the mean ages of the two groups is not statistically significant at the .05 level of confidence.

II. SUBJECTS

Characteristics of the Population

The subjects for this study were all 115 pupils in the five first grade classes at the Weatherbee Elementary School in Hampden, Maine, on February 24, 1969. The mean chronological age of the sixty-one boys and fifty-four girls as of March 1, 1969, was 85.32 months.

The school had 734 pupils in twenty-seven rooms from kindergarten through sixth grade. The school serves the village of Hampden and rural areas nearby.

Hampden, a town of approximately five thousand residents, is located in Central Maine adjacent to Bangor, a city of approximately forty thousand persons.

In a recent economic base study of Hampden (Crawford, 1969), 725 persons of the estimated 1,400 person work force were surveyed by interview and questionnaire. It was found that 21 per cent of the persons contacted were employed within Hampden and 79 per cent outside the town.

primarily in nearby Bangor. Of the workers surveyed, 4 per cent held advanced college degrees, 20 per cent were college graduates at the bachelor's level, 53 per cent were high school graduates, 12 per cent attended high school but did not finish, and 11 per cent had a grade school education or less.

On the basis of the survey, Crawford (1969) estimated that 52 per cent of the total work force of Hampden could be classified as skilled workers, 20 per cent as professional, 20 per cent s semi-skilled, and 8 per cent as unskilled. The mean family income in 1968 for Hampden was estimated in the study to be $6,911.00 compared to a median family income of $5,660.00 for the state of Maine and $5,353.00 for the city of Bangor.

Assignment to Groups

For purposes of this study, each pupil was randomly assigned to one of four sections without regard to homeroom membership. Boys and girls were assigned separately in order to assure the same ratio of boys to girls in each section.

The assignment was begun by numbering all of the boys on a roster from 01 through 61. The girls were numbered from 01 through 54. Next, a table of random numbers (Fisher & Yates, 1963) was opened at random and beginning at a randomly chosen spot on the page, groups of three digits were read from left to right until a number corresponding to one of the page numbers in the table appeared. Then, beginning at the top left hand corner of this new page, pairs of digits were read from left to right until a number between 01 and 61 appeared. The subject whose number corresponded to the number that appeared was assigned to Section A.

The next pupil whose number appeared in the table was assigned to Section B; the next to Section C; and so on until all the boys had been assigned. The same procedure was used to assign all the girls at random to the four sections.

After the assignment of pupils to each section had been completed, it was necessary to randomly select two of the sections to receive the experimental treatment. The sections were numbered one through four and a new starting place in the table of random numbers was found in the same manner as described on the preceding page. The first numbers between one and four to appear randomly on the new page were three and two. Sections C and B, therefore, were designated the experimental group sections and Sections A and D, the control group sections.

III. TREATMENTS

Experimental Group Treatment

The experimental group in this study used Visual Tracking (Geake & Smith, 1962a), a published program designed to develop perceptual skills in reading. Each page of Visual Tracking contains from two to six exercises composed of nonsense words arranged in sentence form like the following example:

> Peant derna nyws torrnib jerst wuttals plagter. Murs bis prid tolik rige mecol, filt rsd turosp. Ness kgzh wuvvop irp ponur zoss paftor pox kotall voxabbntuv quog spon kuab lortn. Wonbe saxy. Dopur hoss ropin vuett poxt turj wanop beryu tolp. Drep celp bec. . . .

Directions for completion of the exercises as printed in the front of the program are:

> This is a book of exercises in which you are to find the letters of the alphabet and draw a line through them (a̸). Each line of make-believe words contains some letters which go together in the alphabet. You are to begin with the first letter "a" in the first line and draw a line through it (a̸). Then find the first "b" after the "a" which you have drawn the line through and draw the line through the "b", too. Then find the first "c" after the "b", then the first "d", and keep on going like that. When you come to the end of the alphabet, write down how long it took. Then begin all over with the next paragraph. Every line has some letters that you will need. If you go through a whole line without finding any letters that you need, then you know that you have made a mistake. If that happens, go back and find your mistake.

The alphabet is printed across the top of each page to aid pupils who are not familiar with all the letters or with alphabetical order.

The _Visual Tracking_ exercises are so arranged that a response to each letter requires discrimination of it from similar letters nearby. Difficulty of the exercises is increased in nine steps by gradually reducing: (1) size of type from large (18 point) to small (10 point), (2) size of letter spacing from wide (4-7 point leading) to narrow (0-2 point leading), and (3) size of line spacing from wide (10 point leading) to narrow (1 point leading). See Appendix A for sample exercises from _Visual Tracking_ illustrating the nine steps of reduction in type size and letter and line spacing.

Control Group Treatment

The instruction provided for the two sections of the control group consisted of teacher-directed games and activities designed to improve listening skills. The lessons for these activities were planned by the investigator and taught by the first grade teacher supervising each control group section. See Appendix B for sample lesson plans.

IV. PROCEDURES IN TREATMENT GROUPS

Beginning February 24, 1969, the experimental group pupils worked on Visual Tracking exercises for fifteen minutes each day for a period of twelve weeks. After three days of instruction and practice, the pupils worked independently at their own rates and with a minimum of assistance from the first grade teachers who supervised the group. During the same period each day, the control group pupils participated in directed listening activities.

In order to facilitate teacher supervision and to avoid contact between the experimental and control groups, all pupils were regrouped daily for the training session. For example, Section A was randomly chosen to receive the control group treatment, so all pupils assigned to that section moved from their respective homerooms to Room 1 for listening activities. The same regrouping procedure was followed for the pupils assigned to the other three sections. The diagram in Figure 3 illustrates the placement of pupils during the daily fifteen-minute training session.

In order to equalize any possible effects of teacher supervision, teachers were rotated among the rooms for the training period according to the schedule shown in Figure 4.

Room 1	Room 2	Room 3	Room 4
Section A 29 pupils (Control)	Section D 28 pupils (Control)	Section B 29 pupils (Experimental)	Section C 29 pupils (Experimental)

Figure 3. Placement of the four sections during the daily training period.

During the eleventh and twelfth weeks, each teacher supervised each group for two days and assisted for two days. For the rest of the school day, pupils remained with their regular homeroom teachers to participate in the usual school activities.

	WEEK -	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
TEACHER	A	1	2	3	4	X[a]	1	2	3	4	X
	B	2	3	4	X	1	2	3	4	X	1
	C	3	4	X	1	2	3	4	X	1	2
	D	4	X	1	2	3	4	X	1	2	3
	E	X	1	2	3	4	X	1	2	3	4

[a]X - Assist with timing in Room 4

Figure 4. Room assignments for teacher supervision during the daily training period.

Instructions to Pupils

On Friday, February 21, 1969, the regular classroom teachers told their pupils that the following Monday they would begin a new procedure. Although each teacher explained the procedure in her own way, each was instructed to emphasize the following points:

1. The pupils were going to have some contact with each of the five first grade teachers.

2. They were going to be in different groups with some pupils from their own homeroom and with some pupils from each of the other homerooms.

3. The pupils were to be regrouped for only fifteen minutes each day and then returned to their regular classes.

4. Different groups were going to do different things. All the things would be new and interesting. The teachers did not mention the investigator, a research project, or the University of Maine. Rather, the project was presented as a new experience devised by the teachers in the fulfillment of their roles.

After the project had been described, the teachers answered the pupils' questions. Then they passed out color-coded name tags used both to identify the pupils for the other teachers and to insure that pupils arrived at the correct room each day. The doors of the four rooms used for the treatments were marked with the tag color of the group meeting there.

Finally, teachers and pupils practiced the regrouping to acquaint the pupils with the procedure for movement and to solve any procedural problems before the actual experiment began.

The training sessions were held each day following the children's lunch period. As soon as the pupils returned to their homerooms from lunch recess, they picked up their name tags and moved to the proper rooms for the experimental or control treatments. Every aspect of the regrouping procedure went smoothly on the practice day and throughout the duration of the experiment.

Experimental Group Procedure

On February 24, 1969, the teachers supervising the experimental groups presented orally the directions for completing the exercises in Visual Tracking and then demonstrated the procedure with an illustrative page on an overhead projector. Through the use of teacher demonstrations and mimeographed practice exercises for each pupil, every effort was made to be certain that all understood the directions.

After the practice exercises had been used for three days, the pupils were ready to begin working in the Visual Tracking program. By the end of the first week, nearly all pupils were able to complete the exercises in the program with a minimum of assistance. By the end of the second week, all pupils were able to work independently.

After the second week, some pupils needed occasional help in locating a letter or in correcting a problem caused by marking the wrong letter. These minor problems decreased in frequency throughout the term of the experiment but never entirely disappeared; a minimum of teacher assistance was always needed.

Timing procedure. In order to motivate the pupils to work more rapidly on the tracking, the completion time for every exercise was computed and plotted on a line graph progress chart at the end of each section in each pupil's Visual Tracking program. According to the authors (Geake & Smith, 1962b, p. 2), completion times are determined in order to give pupils immediate knowledge of results and are recorded on the pupils' progress charts for their reward value. See Appendix C for a sample progress chart.

Completion times were computed through the use of numerals on the chalkboard that were changed at thirty second intervals. The numerals merely functioned as indicators of time intervals and held no other significance.

When a pupil was ready to begin a tracking exercise, he recorded, below the exercise, the number showing on the chalkboard. Upon completing a paragraph, the pupil again recorded the number showing on the chalkboard and then raised his hand to signal the supervising teacher. The teacher recorded the elapsed time beside the completed exercise and then encouraged the child to attempt to improve his time on the next exercise. Teacher aides were hired to transfer the completion times to the pupils' progress charts.

Completion procedure. Since the experimental group pupils worked independently at their own rates, they varied considerably in the time they took to complete the program. The first pupil to finish completed the program during the ninth week. Each day thereafter, a few additional subjects finished the program until, at the end of the twelve week experiment, 91 per cent of the children had finished the *Visual Tracking* program. Of the ten pupils who did not complete the program, five were working in Part III, two in Part IV, and three in Part V, the final part of the program.

In order to provide continued tracking practice, the pupils who completed the Visual Tracking program early were given mimeographed sets of tracking exercises and were shown how to plot their completion times on a teacher-prepared graph. For the remainder of the experiment, these pupils worked on the mimeographed exercises and plotted their completion times as computed by the teacher.

Control Group Procedure

For the control group pupils, participation in the listening activities involved much aural and visual attention, some verbal response, and occasional physical movement. Care was taken to avoid any visual or visual-motor activities that resembled the behavior required by the visual tracking training.

In order to avoid creating the impression that control sections were having an extra recess while experimental sections were required to do hard work, teachers attempted to balance the appeal of the experimental and control activities. The teachers tried to keep the listening activities pleasant for the pupils while avoiding an over-emphasis upon frivolity and play.

The appeal of the visual tracking task was enhanced by the ease of obtaining success, the continual feedback from the program itself, the supervising teachers' indications of success, the stimulation of competing with oneself to improve completion times, and by the de-emphasis on mistakes or failure.

The rotation of teachers also avoided an imbalance in the appeal of the activities. Because each teacher supervised each group for an equal time, no one teacher could become associated with any one group and thereby affect the attractiveness of that group's activity.

V. PILOT STUDY

Pilot Study Procedure

A pilot study utilizing the Reading Eye was conducted in May, 1969, with a group of forty-eight first grade pupils at the Asa Adams Elementary School in Orono, Maine. The pilot study was designed to facilitate the data collection in the main study in several ways.

Experience for examiners. Both the investigator and another doctoral student who assisted with the testing in the main study gained valuable experience in photographing the eye movements of first grade children and in scoring the eye-movement films.

Procedural adaptations. Directions and procedures for testing first grade subjects with the Reading Eye were developed in the pilot study. Although the directions in the Reading Eye manual (S. E. Taylor, 1960) were followed, it was necessary to simplify the directions to the pupils and to take extra precautions to avoid confusion or concern about the testing. The directions and procedures developed are described in detail on pages 94-98.

As a result of a difficulty that arose in scoring the eye-movement films during the pilot study, one addition was made to the recommended testing procedure. In the Reading Eye manual (S. E. Taylor, 1960, p. 25), directions for adjusting the camera to the subject were followed by these instructions:

> Then close the aperture door, starting the filming action. Point to the upper right-hand o, then the lower right o, then the upper left o, then the lower left o, meanwhile saying, "Look at this o, this o, this o, and this o." Then drop the target card, . . .

Although the manual does not explain the reason for having the subject look at the o's on the corners of the target card before reading, presumably it serves to get the subject's eyes in motion after a period of fixating upon an X on the target card. See Appendix D for a photograph of the Reading Eye including the target card. Also, the procedure should, as the manual states on page 34, provide an eye-movement record that is ". . . always characterized at the beginning by the long fixations made while the examiner said, 'Look at this o, this o, this o, and this o.' Immediately following this, the reader will swing into the first line of print or staircase." (S. E. Taylor, 1960, p. 34).

When scoring films during the pilot study, however, it was discovered that many of the pupils had not fixated directly upon the o's, but had moved their eyes across the card creating a pattern on the film much like the pattern created by actual reading. In those cases, it was impossible to determine with certainty which lines at the beginning of the film record were associated with "looking at the o's" and which represented the movements of the eyes while reading the first line.

In order to avoid the possibility that the eye-movement pattern corresponding to "looking at the o's" would be mistaken for the eye-movement pattern corresponding to actual reading of the first line of print, an additional procedure was developed and tested during the pilot study. The change consisted of asking a pupil to close his eyes after looking at the third o. Then, after a one-second pause, the pupil was asked to reopen his eyes and look at the last o on the target card.

When the pupil closed his eyes for one second, the beam of light shining on his corneas was interrupted and a corresponding break in the eye-movement line appeared on the developed film. Then it was a simple matter to locate the break in the line to identify the point at which the pupil looked at the last o before reading the test selection.

Equipment adaptations. The following minor changes were gradually made to equipment during the pilot study as they were needed to improve the results obtained:

1. Pictures of animals were placed over the X on the target card at which the pupil looked while the examiner focused the camera. Instead of following the exact statement in the manual (S. E. Taylor, 1960, p. 24), ". . . Hold your head as still as possible and look at the X in the center of the card," the examiner in this study said, ". . . look at the [duck or mouse] on this card." It was found that the animal held the attention of first grade pupils for a longer time than did the X. The examiner, therefore, had additional time in which to complete the focusing. See Appendix E for a photograph of

a target card with an animal in place of the X.

2. Firm sponge rubber pads were attached to the inside of both head steadiers on the Reading Eye. These pads improved the effectiveness of the steadiers in holding the head in place. In addition, the pupils reported that the steadiers felt more comfortable with the pads attached. See Appendix D for a photograph of the Reading Eye with sponge rubber pads attached.

3. A wooden school desk was lowered at the front by cutting a two-inch piece off each front leg. This modified table provided a slant for the Reading Eye that enabled the first graders to view the reading selection at a more comfortable angle than when a flat table was used. This adaptation was necessary because most of the pupils were so short that the adjustment needed for proper viewing exceeded the adjustment capabilities of the Reading Eye on a flat surface. See Appendix F for photographs of a first grade pupil using the Reading Eye at a flat table and at a modified table.

Easier test selections. Four easier test selections and accompanying comprehension questions written by the investigator were used in the pilot study along with the regular Reading Eye test selections. These easier selections were designed to serve as alternate test selections in the main study for use with pupils unable to read the regular Reading Eye first grade test selections with adequate comprehension. See Appendix G for the complete texts of the easier test selections and their accompanying comprehension questions.

The writing and printing of the new selections was carefully designed to make them as nearly identical to the regular Reading Eye selections as possible except for vocabulary. See Figures 5a and 5b for a comparison of a Reading Eye first grade test selection with an easier test selection.

The vocabulary list used to write the easier test selections was developed from the vocabularies of three primer level reading textbooks. One primer, _Dot and Jim_ (Harris, Creekmore, & Greenman, 1964), had been used in all of the first grades at Weatherbee School. In addition, most of the first grade pupils at Weatherbee School and all at Adams School had read one of the other two primers, _Fun With Our Friends_ (Robinson, Monroe, & Artley, 1962) and _The Little White House_ (Russell & Ousley, 1961). Only words that appeared both in _Dot and Jim_ and one of the other two primer textbooks were used in writing the four

John had a red hen for a pet.
He took his hen to school one day.
The children at school laughed at John.
"What good is a hen," they said.
"You cannot play with a red hen."
"My pet hen lays eggs," said John.
"She lays one brown egg every day.
The next morning I eat the egg.
That is why she is my pet."

Figure 5a. One of the Reading Eye first grade level test selections. (Reprinted by permission of Educational Developmental Laboratories, Inc.)

Jim walked to the dog house.
He saw a little dog in it.
The little dog was white and black.
Jim said, "You are not my dog."
He ran fast to get his mother.
She came to see the new dog.
It was not there. Where did it go?
They looked and looked for the dog.
But they did not see it.

Figure 5b. One of the easier test selections written by the investigator for use with the Reading Eye.

easy selections.

The following steps were taken to insure comparability of test selections:

1. The countable lines of the new selections were made equal to the countable lines of the Reading Eye selections in number of words and actual measured length.

2. Each story concerned a child and his or her pet; two concerned boys and two concerned girls. The Reading Eye selections followed the same pattern.

3. The ten comprehension questions were written to require recall of detail just as the ten comprehension questions for the regular Reading Eye selections require.

4. The Reading Eye selections were printed with black ink on white card stock with 13.7 point Caledonia type with 4 point leading between lines and variable leading between letters and words to equalize line lengths. The new selections depart only in that 14 point Baskerville type was used because 13.7 point Caledonia type was not available.

The slight differences between types used for the Reading Eye selections and the new easier selections are judged to be so insignificant that they would have no effect upon reading performance. The Caledonia and Baskerville types are both serif letters and both belong to the family commonly known as Roman type. In summarizing studies of legibility of various type sizes and styles, Tinker (1965, p. 134) concluded:

> Therefore, it seems safe to assume that currently used typefaces, all printed in the same point size, leading, line width, and paper stock, would be read with approximately the same speed (be equally legible) but that readers would rate some to be more legible than others.

<u>Readability of the test selections</u>. According to S. E. Taylor (1960, p. 17), the mean readability of the eight Reading Eye test selections was determined, through application of the Spache readability formula (Spache, 1966, pp. 141-151), to be 1.83, high first grade level. Although the vocabulary for the easier selections was taken from primer level readers, the mean readability of these four selections was determined to be 1.76, also high first grade level.

The similarity in readability levels for the two sets of test selections is due largely to the fact that the Spache formula is more heavily weighted for sentence length than for difficulty of vocabulary. By making the line lengths of the easier selections equal to the line lengths of the Reading Eye selections, as described on page 88, the readability levels were nearly equated despite some vocabulary differences. It was expected, however, that the investigator's easier selections would, in fact, be easier for the first graders to read because the vocabulary for these selections was taken directly from reading books they had used in school and because the words contain fewer letters on the average than the words in the Reading Eye selections.

Pilot Study Findings

In the pilot study, the twenty-three pupils who could read the Reading Eye test selections with adequate comprehension were photographed while reading two Reading Eye selections and two easier selections. For these pupils, the order of presentation was rotated so that approximately one-half read the Reading Eye selections first and the other half read the easier selections first. The twenty-five remaining pupils read two of the easier test selections.

Many films were rendered unscorable because of excessive head movements and spoiled film. These problems were especially prevalent during the first two days of testing as the examiners gained experience and worked to improve the testing procedures and equipment. Because time was short, no retests were scheduled in cases of unscorable films; consequently, from zero to four sets of eye-movement scores were available for each pupil.

Analysis of forty-three pairs of pilot study test scores was undertaken to determine, for those pupils who read both reading selections, if there were significant differences between mean eye-movement scores based upon reading each test selection. Accordingly, for each eye-movement measure, differences between mean scores obtained as pupils read the Reading Eye selections and mean scores obtained as pupils read the easier selections were tested for statistical significance by a t-test for correlated groups. Results of these t-tests are presented in Table II.

As revealed in Table II, t-tests indicated that mean performances on the Reading Eye and easier test selections differed significantly at the .05 level of confidence for each eye-movement measure. Therefore, it was judged probable that a pupil in the main study would not obtain equivalent or nearly equivalent eye-movement scores when reading either a Reading Eye selection or an easier selection. It was not considered possible, therefore, to combine all scores for each group in the main study without regard for the reading selections read. Instead, it was necessary to decide among three alternatives: (1) to use the Reading Eye test selections for all pupils in the main study, (2) to use the easier test

TABLE II

t-TESTS FOR CORRELATED GROUPS READING EYE VERSUS EASIER SELECTIONS FROM PILOT STUDY

N=43

Selection	$\bar{X}$	S.D.	$S.E._{\bar{X}}$	t
	Number of Fixations			
Reading Eye	190.93	36.18	4.78	3.27[a]
Easier	175.30	25.5		
	Number of Regressions			
Reading Eye	31.16	14.34	2.30	2.41[a]
Easier	26.23	10.55		
	Average Span of Recognition			
Reading Eye	.5414	.1067	.0145	2.86[a]
Easier	.5827	.0845		
	Average Duration of Fixation			
Reading Eye	.3353	.0546	.0058	2.66[a]
Easier	.3199	.0351		
	Rate with Comprehension			
Reading Eye	99.46	26.14	2.88	3.52[a]
Easier	109.63	21.41		

[a]$p < .05$

selections for all pupils in the main study, or (3) to use both selections and devise a method of equating the obtained scores.

The first alternative was rejected because it was probable that a large number of pupils would have been disqualified from the study because of an inability to read the Reading Eye selections with adequate comprehension. The second alternative was rejected because of the decision to gather test-retest reliability data for eye-movement scores based upon reading the Reading Eye test selections.

The third alternative was, therefore, implemented by drawing lines of relationship between the eye-movement scores obtained while pilot study pupils read the Reading Eye selections and those obtained while the same pupils read the easier selections. These lines of relationship or "equi-percentile curves" (Flanagan, 1951, pp. 752-755) were computed to equate the Reading Eye selection and easier selection scores for number of fixations, number of regressions, and rate.[2]

In the main study, fixation, regression, and rate

[2]Since the average span of recognition and average duration of fixation scores are derived from the fixation and rate scores, it was not necessary to compute lines of relationship for span and duration.

scores for each pupil who read an easier selection were transformed into equivalent Reading Eye selection scores, i.e., the comparable scores that would have been expected if the pupil had been able to read a Reading Eye selection with adequate comprehension. Using the transformed scores for fixations and rate, the average span of recognition and average duration of fixation scores were calculated.

VI. DATA COLLECTION AND ANALYSIS--MAIN STUDY

Testing Procedure

The first grade pupils at Weatherbee School were introduced to the Reading Eye by the researcher and an assistant by means of a demonstration in each room before the data collection began. The pupils were told that the experimenters were students from the University of Maine who would appreciate help in their efforts to learn to photograph eye movements. The pupils were eager to participate.

At the classroom demonstration, the pupils were told how the Reading Eye functions and were able to see it in operation as the eye movements of the teacher or a classmate were photographed. They were encouraged to ask questions about the camera and the testing procedure so that they might be familiar with the whole process.

Eye movements of all pupils in the study were photographed during the period of May 26, 1969, through June 3, 1969. The testing was conducted in the room available daily for exclusive use by the examiners. The two cameras were placed at opposite ends of the room in a position that prevented pupils from observing each other during the filming. The presence of another pupil being tested in the same room did not appear to affect the test performance of any subject; in fact, the presence of another pupil and examiner seemed to put many pupils at ease.

The classroom teachers sent pupils for testing in random order so that neither examiner knew the experimental or control group status of any pupil whom he tested. When a pupil entered the room, he was asked to sit before one of the cameras. Then the examiner talked informally with the pupil to establish rapport and wrote the pupil's name beside a number on the test record sheet. This number became the pupil's identification number and was later recorded on the film to identify his eye-movement record.

Oral pretest. Before adjusting the camera to the pupil, the examiner handed him one of the regular Reading Eye first grade test selections and asked him to read it aloud. This oral pretest, described in the manual (S. E. Taylor, 1960, p. 20), was given to ascertain that the pupil could successfully read the Reading Eye test selections.

Pupils who began the oral pretest with obvious fluency or with extreme difficulty were not required to finish reading the entire card. Their ability or inability to read the Reading Eye selections was evident. For the pupils who were allowed to read the entire selection orally, performance was judged satisfactory if they made four or fewer oral reading errors. These pupils then read an alternate Reading Eye first grade test selection before the camera. Pupils who made five or more errors on the oral pretest were given one of the easier test selections to read before the camera.

Photographing the eye movements. Once determination of the appropriate test selection had been made, the pupil was asked to move his chair as close to the camera table as possible and to rest his chin upon the camera's chinrest. Then the examiner adjusted the height of the camera, checked to be certain that the pupil's forehead was resting against the headrest pads, and moved the head steadiers into position against the side of the pupil's head. See

Appendix H for an illustration of the use of the Reading Eye. After checking with the pupil to determine that he was comfortable and that he could see the entire target card, the examiner instructed him to look at the animal in the middle of the target card. While the pupil's eyes were fixated upon the animal, the examiner focused the beads of reflected light in the reflex viewing window.

As the examiner was working to focus the reflections, he presented the following instructions to the pupil:

> In a minute I am going to have you read a short story to yourself. Please read the story one time and hold your head as still as you can while you read it. When you are through, I will ask you some questions about the story.

After the focusing was completed, the examiner told the pupil to look at each of the o's on the target card while he checked to be certain that the beads of light stayed within the recording area of the reflex window. When all was in order, the examiner closed the reflex aperture door to start the photography and said:

> Look at this o [examiner points to each o], this o, this o, close your eyes, . . . open them, and look at this o. [As the examiner lowered the target card to expose the reading selection at this point, he continued the directions.] Now read this story about [name of child in story] to yourself one time. Try to remember what you have read and close your eyes when you are through.

During the reading, the examiner watched the pupil for symptoms of head movement, reading distress, or other factors that might influence the quality of the film record or the pupil's comprehension. These signs were noted on the record sheet to aid later interpretation of the film.

When the pupil closed his eyes to indicate that he had finished reading, the examiner stopped the machine, released the head steadiers, and told the pupil to open his eyes and to sit back and relax.

After the pupil was comfortable, the examiner asked him to answer ten yes-no questions as a check of his reading comprehension. According to S. E. Taylor et al. (1960, p. 9):

> The quiz was not designed to provide an exact measure of the degree of comprehension, but rather to give a general indication of whether or not the subject was reading.

Since this study was concerned with measures of eye movements during silent reading, it was important to know that each pupil had actually read a selection.

After the questions had been answered, the examiner made a positive remark such as "fine" or "good job." Then he recorded the pupil's comprehension percentage score and the number and level of the selection read on the record sheet. He also identified the pupil's film by inserting in the camera a plastic strip on which he had written the pupil's identification number, selection number, level, and comprehension score. Subsequent readings, if needed, were identified with the same number plus a letter. For example, film for the original test and two subsequent retests for Pupil Number 28 would have been identified as: "28," "28a," and "28b."

If no retest was necessary, the pupil returned to his room. If, however, a retest was needed for reliability computation or because of low comprehension on the first reading, the same testing procedure was followed with another reading selection except that the oral pretest was omitted.

Retest procedure. Pupils were retested with an alternate test selection for one or more of the following reasons: (1) because comprehension was below 70 per cent on the first reading, (2) because another set of measures was needed in order to compute reliability coefficients, or (3) because a film obtained the previous day was found to be unscorable after it had been developed.

When a pupil's score on the comprehension check for a selection was below 70 per cent, retesting was conducted immediately. A few pupils who read the Reading Eye selection well enough on the oral pretest to be tested with that level before the camera had extremely low comprehension scores after the filming. In these cases, the retest was performed with one of the easier stories instead of an alternate Reading Eye selection in order to avoid the possibility of continued failure to comprehend. In all other cases, retesting was conducted with an alternate test selection of the same level used for the initial test.

All pupils who read a Reading Eye selection with low comprehension for their first eye-movement photograph eventually achieved a comprehension score of 70 per cent or better after reading an equivalent Reading Eye selection or an easier selection on the first or second retest.

Among the pupils who read one of the easier selections for their first eye-movement photograph were seventeen pupils who scored 50 per cent or below on the comprehension check and who evidenced extreme difficulty with the reading. After the comprehension check, these pupils were asked to read the same story orally to the examiner. Nine of these pupils were able to read only a few words so further testing was discontinued and they were dropped from the study. Three other pupils were dropped from the study after failing to achieve a comprehension score of 70 per cent on two retests. The remainder were able to achieve an adequate comprehension score on subsequent retests.

In order to provide an additional set of eye-movement scores for computation of reliability coefficients, an attempt was made to retest immediately every second pupil with an alternate reading selection. In some cases it was not possible to maintain the every-other-pupil pattern because some pupils scheduled for retest were able to read only one selection with adequate comprehension. In those cases, the following pupil was retested to obtain the reliability data. In a few cases the extra films taken for this purpose were found to be unscorable after all the testing had been completed. Therefore, two sets of scores are available for only forty-four pupils instead of forty-eight pupils.

Retests were arranged the following day for pupils whose eye-movement photographs were found to be unscorable. The primary cause of unscorable films was excessive head movement. When head movement was excessive, the eye-movement lines on the film contained so many curves that it was impossible to count fixations. Sometimes the head was moved so far that light reflected from the pupil's eyes either did not reach the film or reached it out of focus. In a very few cases, portions of film were ruined when an examiner accidentally exposed them to light while loading, unloading, or developing them. See Appendix I for a sample film illustrating the effects of head movement upon the eye-movement lines.

No pupils were eliminated from this study because of head movement or poor photography. By structuring the testing procedure to reduce problems of administration, by using the easier test selections, and by retesting when films were poor, it was possible to obtain at least one scorable film for each pupil who could read with adequate comprehension.

Three new pupils entered Weatherbee School during the fifth week and, therefore, could not be included in the experiment. To put them at ease, they were placed with the control group pupils during the daily training sessions. Their eye movements were photographed, but their test scores were not included in the statistical analyses.

At the conclusion of the study, the researcher asked the first grade teachers to respond in writing to the following written request:

> Please list any reactions or observations you have regarding the visual tracking and listening activities. Make any comments you wish that will help me [the

researcher] to describe the value or lack of value of the experiment for your pupils.

Scoring the photographs. All eye-movement films were developed and scored by the investigator in accordance with directions contained in the Reading Eye manual (S. E. Taylor, 1960). In addition, an assistant rescored the primary film record for each of the ninety-six pupils in the study in order to provide an additional set of data for determination of inter-rater reliability. The retest film records to be used for determination of test-retest reliability were scored only by the researcher.

In order to avoid contamination, the investigator and the other examiner erased all grease pencil marks from each film record and used separate scoring sheets. The scoring sheets contained only the pupil's identification number, the level and number of the selection read, and the comprehension score for each film to be scored.

Data Analysis

After all scoring had been completed, pupils' names were added to the score sheets so that the test scores could be grouped according to sex and experimental or control group assignment.

Before the statistical analysis could be performed by computer, it was necessary to equalize the number of male and female pupils in the experimental and control groups. Table III presents the membership status of each group by sex from the beginning of the study until the end. It can be seen by reference to Table III that it was necessary to randomly exclude the data for two experimental group girls and three control group boys in order to balance the groups.

This random exclusion of pupils was begun by numbering all of the girls in the experimental group for whom test data were available from 01 through 26. The boys in the control group for whom test data were available were numbered from 01 through 27. Then, a page was found at random in a table of random numbers (Fisher & Yates, 1963) by the same method used when making the initial assignments as described on page 71. Next, beginning at the upper left hand column, pairs of digits were read horizontally until two numbers between 01 and 26 appeared. The data for the two experimental group girls whose numbers appeared first in the table were excluded from the study. The same procedure, including the random selection of a new page in the table of random numbers, was followed to exclude the data for three control group boys.

TABLE III

EXPERIMENTAL AND CONTROL GROUP MEMBERSHIP
ADJUSTMENTS AND TOTALS

MEMBERSHIP ADJUSTMENT	TOTAL	EXPERIMENTAL		CONTROL	
		Boys	Girls	Boys	Girls
Initial Group Assignment	115	30	28	31	26
Left school - moved	-2	0	0	-1	-1
New pupils	[3]	0	0	0	0
Eliminated - No test scores due to inability to read any selection.	-12	-6	-2	-3	-1
Membership after films were scored.	101	24	26	27	24
Eliminated at random to balance groups.	-5	0	-2	-3	0
Final Membership	96	24	24	24	24

After the groups had been balanced, the eye-movement data were punched on IBM cards and sent to the University of Maine Computing Center where statistical analysis was performed on an IBM 360 computer. A factorial analysis of variance program for randomized groups and a linear correlation program were used for analysis of the data.

CHAPTER IV

ANALYSIS OF THE DATA

This chapter presents the data that were gathered to test the three hypotheses proposed in Chapter I, to determine test-retest and inter-rater reliabilities of the eye-movement measures, and to make comparisons with Reading Eye norms. Available for analysis were eye-movement scores for all ninety-six subjects; retest eye-movement scores for forty-four subjects; and, for the ninety-six subjects, eye-movement scores based upon a separate scoring of the films by an assistant.

I. THE EFFECTS OF VISUAL TRACKING TRAINING

In order to determine, for a selected population of first grade pupils, the effects of visual tracking training upon five selected components of silent reading performance measured by eye-movement photography, the following hypotheses were tested for statistical significance by application of factorial analysis of variance to the data:

1. There will be no significant differences between mean performances of the experimental and control group subjects on the following eye-movement measures: (a) number of fixations, (b) number of regressions, (c) average span of recognition, (d) average duration of fixation, and (e) rate with comprehension.

2. There will be no significant differences between mean performances of subjects of either sex on any of the five eye-movement measures.

3. There will be no significant interaction between sex and visual tracking training on any of the five eye-movement measures.

Table IV presents mean eye-movement scores for each group in this study. According to Table IV, performance of the experimental group pupils was slightly superior to the control group pupils on all measures but none of the differences was statistically significant.[3]

The mean number of fixations was 232.66 for the experimental group and 239.95 for the control group; a difference of 7.29 fixations. The mean number of

---

[3]For number of fixations, number of regressions, and average duration of fixation, a lower score represents better reading performance. For average span of recognition and rate with comprehension, better reading performance is represented by a higher score.

TABLE IV

MEAN SCORES ON THE FIVE
EYE-MOVEMENT MEASURES
FOR EACH GROUP

GROUP	N	MEAN SCORE				
		Fixations[a]	Regressions[a]	Span	Duration	Rate
Treatment (A)						
Experimental	48	232.66	38.66	.4633	.5124	67.36
Control	48	239.95	43.74	.4567	.6209	65.22
Sex (B)						
Boys	48	255.78[b]	46.87	.4399	.6886	58.46[b]
Girls	48	216.84[b]	35.52	.4801	.4447	74.12[b]
Interaction (AXB)						
Exp. Boys	24	255.20	45.58	.4289	.5879	58.49
Exp. Girls	24	210.12	31.73	.4977	.4369	76.24
Cont. Boys	24	256.35	48.17	.4508	.7892	58.42
Cont. Girls	24	223.56	39.31	.4625	.4526	72.01
All Subjects	96	233.43	40.54	.4562	.5612	65.86

[a]In accordance with the Reading Eye manual, these scores represent number of fixations and number of regressions per one hundred words. Since first grade pupils read only fifty words before the camera, the obtained number of fixations and regressions for each subject is doubled to base the scores upon one hundred words read (S. E. Taylor, 1960, pp. 37-39).

[b]Differences significant at the .05 level. See Tables V and IX.

regressions was 38.66 for the experimental group and 43.74 for the control group; a difference of 5.08 regressions. The mean span of recognition was .4633 words for the experimental group and .4567 words for the control group; a difference of .0066 words. The mean duration of fixation was .5124 seconds for the experimental group and .6209 seconds for the control group; a difference of .1085 seconds. The mean rate of reading was 67.36 words per minute for the experimental group and 65.22 words per minute for the control group; a difference of 2.14 words per minute.

Tables V through IX present the analysis of variance for each of the five eye-movement measures. Values of F significant at the .05 level of confidence were required to reject the null hypotheses.

Reference to Tables V through IX reveals that differences between experimental and control group means for each eye-movement measure failed to reach the required level of significance. The first hypothesis, therefore, was not rejected. Visual tracking training did not produce significantly better reading performance, as measured by eye-movement photography, in the experimental group than in the control group.

According to Table IV, differences between mean eye-movement scores of the boys and girls were statistically significant for two measures. The mean number of fixations was 255.78 for the boys and 216.84 for the girls, a difference of 38.94 fixations. According to Table V, this difference is statistically significant at the .05 level of confidence. The mean rate of reading with comprehension was 58.46 words per minute for the boys and 74.12 words per minute for the girls, a difference of 15.66 words per minute. According to Table IX, this difference is statistically significant at the .05 level of confidence.

TABLE V

ANALYSIS OF VARIANCE FOR
NUMBER OF FIXATIONS

Source	df	Sum of Squares	Mean Square	F
Treatment (A)	1	1,276.02	1,276.02	.18
Sex (B)	1	36,386.90	36,386.90	5.20*
Interaction (AXB)	1	906.53	906.53	.13
Error	92	644,176.00	7,001.91	
Total	95	682,745.45		

*$p < .05$

TABLE VI

ANALYSIS OF VARIANCE FOR
NUMBER OF REGRESSIONS

Source	df	Sum of Squares	Mean Square	F
Treatment (A)	1	620.17	620.17	.66
Sex (B)	1	3,094.00	3,094.00	3.27
Interaction (AXB)	1	150.00	150.00	.16
Error	92	87,003.50	945.69	
Total	95	90,867.67		

TABLE VII

ANALYSIS OF VARIANCE FOR AVERAGE
SPAN OF RECOGNITION

Source	df	Sum of Squares	Mean Square	F
Treatment (A)	1	.001	.001	.07
Sex (B)	1	.04	.04	2.66
Interaction (AXB)	1	.02	.02	1.34
Error	92	1.34	.01	
Total	95	1.401		

TABLE VIII

ANALYSIS OF VARIANCE FOR AVERAGE DURATION OF FIXATION

Source	df	Sum of Squares	Mean Square	F
Treatment (A)	1	.28	.28	.48
Sex (B)	1	1.43	1.43	2.43
Interaction (AXB)	1	.21	.21	.35
Error	92	53.90	.59	
Total	95	55.82		

TABLE IX

ANALYSIS OF VARIANCE FOR RATE WITH COMPREHENSION

Source	df	Sum of Squares	Mean Square	F
Treatment (A)	1	110.17	110.17	.11
Sex (B)	1	5,892.21	5,892.21	5.98*
Interaction (AXB)	1	103.83	103.83	.10
Error	92	90,639.60	985.21	
Total	95	96,745.81		

*$p < .05$

The mean number of regressions was 46.87 for the boys and 35.52 for the girls, a difference of 11.35 regressions. According to Table VI, this difference is not statistically significant.

The mean span of recognition for the boys was .4399 words and for the girls was .4801 words, a difference of .0402 words. According to Table VII, this difference is not statistically significant.

The mean duration of fixation for the boys was .6886 seconds and for the girls was .4447 seconds, a difference of .2439 seconds. According to Table VIII, this difference is not statistically significant.

In accordance with the findings reported above for the sex variable, the second hypothesis was rejected for number of fixations and for rate with comprehension. For number of regressions, average span of recognition, and average duration of fixation, the second hypothesis was not rejected.

Tables V through IX reveal that no significant interaction was found between sex and visual tracking training. A significant interaction F ratio would have indicated that the treatment variable was interacting with the sex variable to produce higher reading performance for pupils of one sex and lower performance for pupils of the other sex. Since no such significant interaction was found in this study, the third hypothesis was not rejected.

Summary

Analysis of the data presented in this section revealed no significant differences on the treatment variable or on the interaction of treatment and sex. The first and third hypothesis, therefore, were not rejected.

The second hypothesis was rejected for number of fixations and rate with comprehension because significant F ratios were found on the sex variable for these two eye-movement measures. The second hypothesis was not rejected for number of regressions, average span of recognition, and average duration of fixation because values for F for these eye-movement measures were non-significant.

In this study, effects of visual tracking training yielded neither significantly higher eye-movement scores for the experimental group nor significant interaction between the variables. On the sex variable, however, the girls were found to read the test selections at a significantly more rapid rate and with significantly fewer fixations than the boys while maintaining an acceptable level of comprehension.

II. TEST-RETEST RELIABILITY

Eye-movement scores were available for determination of test-retest reliability from forty-four pupils in the study. Retest scores for thirty of the pupils were obtained at the same sitting. For the remainder, it was necessary to obtain a retest film the following day because one of the films from the first sitting was found to be unscorable.

Table X reports the test-retest reliability coefficients obtained from application of a linear correlation program to the data.

TABLE X

TEST-RETEST RELIABILITY COEFFICIENTS
FOR EYE-MOVEMENT MEASURES

N=44

EYE-MOVEMENT MEASURE	r
Number of Fixations	.88
Number of Regressions	.84
Average Span of Recognition	.68
Average Duration of Fixation	.66
Rate with Comprehension	.80

Substantial consistency of test performance is indicated by reliability coefficients of .88 for number of fixations, .84 for number of regressions, and .80 for rate with comprehension. Moderate consistency is indicated by reliability coefficients of .68 for average span of recognition and .66 for average duration of fixation.

Inasmuch as no previous study has reported test-retest reliability coefficients for eye-movement scores of first grade pupils, coefficients obtained in this study can be compared only with those from studies of older subjects.

Reliability coefficients from earlier studies reported in Table I, page 55, ranged from .61 to .92 for number of fixations, from .58 to .93 for number of regressions, and from .52 to .96 for rate of reading. For each of these eye-movement measures, the reliability coefficients obtained in this current study are among the highest reported.

The reliability coefficients in this study are higher than most coefficients reported in earlier studies using elementary school pupils as subjects. For example, Table I, page 55, indicates that with fourth and fifth graders Eurich (1933b) obtained reliability coefficients of .70 to .87 for fixations, .68 to .87 for regressions, and .62 to .69 for duration. Also, with fourth through sixth graders, Broom (1940) obtained reliability coefficients of .79 for fixations, .76 for regressions, and .78 for rate.

Because reliability coefficients for average span of recognition and average duration of fixation are relatively meaningless, no previous study has reported a coefficient for span of recognition and only a few studies have reported coefficients for average duration of fixation. Average span of recognition is determined by dividing the number of words read, a constant, by the number of fixations. Average duration of fixation is derived from the number of fixations and rate of reading scores. With the Reading Eye neither span of recognition nor duration of fixation is measured directly from the eye-movement film. Both span and duration vary according to variations in one or more of the other eye-movement measures; therefore, the consistency of the span and duration scores is dependent upon the consistency of the other measures.

In Chapter II, it was indicated that collection of test-retest reliability data was needed for two reasons: (1) it was not certain that the fifty-word test selections were long enough to yield reliable eye-movement scores, and (2) no previous study had reported reliability coefficients for the eye-movements of first grade pupils. The substantial reliability coefficients reported in Table X: .88 for number of fixations, .84 for number of regressions, and .80 for rate with comprehension, provide evidence that the fifty-word test selections are long enough to yield reliable eye-movement scores with first grade pupils. The obtained reliability coefficients are of sufficient magnitude to support the conclusion of Tinker (1936b, p. 745) and Broom (1940, p. 208) that eye-movement photography has adequate reliability for group measurement and to extend applicability of the conclusion to first grade pupils.

Reading Eye Versus Easier Selections

The reliability coefficients reported in Table X are based upon two readings by forty-four first grade pupils. Twenty-one pupils read Reading Eye test selections; twenty-three read easier test selections. In order to determine whether the reliability of each eye-movement measure differed according to the selection read, separate reliability coefficients were computed for the Reading Eye and easier test selections. These separate reliability coefficients are reported in Table XI.

It may be seen from Table XI that pupils who read the easier selections in this study had much more reliable scores for number of fixations and number of regressions than did those who read the Reading Eye selections. Reliability coefficients of .91 for number of fixations and .90 for number of regressions were obtained for pupils who read the easier test selections. These coefficients are among the highest reported in any study.

TABLE XI

TEST-RETEST RELIABILITY COEFFICIENTS
FOR READING EYE AND EASIER
TEST SELECTIONS

	N=21	N=23
EYE-MOVEMENT MEASURE	Reading Eye Test Selections	Easier Test Selections
Number of Fixations	.51	.91
Number of Regressions	.36	.90
Average Span of Recognition	.64	.66
Average Duration of Fixation	.80	.65
Rate with Comprehension	.81	.70

On the other hand, reliability coefficients of .51 for number of fixations and .36 for number of regressions were obtained for pupils who read the Reading Eye test selections. These coefficients are among the lowest reported in any study.

Again, the reliability coefficients for average span of recognition and average duration of fixati are dependent upon variations in number of fixations and rate and have little meaning as measures of consistency of reading performance. For rate with comprehension, pupils read the Reading Eye test selections with more consistency of performance than those who read the easier selections. The reliability coefficient of .81 for rate on the Reading Eye selections indicates substantial consistency while the lower reliability coefficient of .70 for rate on the easier selections indicates moderate consistency of performance.

Summary

The test-retest reliability coefficients reported in this study were of sufficient magnitude to conclude that eye-movement photography has adequate reliability for group measurement of first grade pupils who read fifty-word test selections. When the test-retest data were divided according to selections read, it was found that pupils read the easier test selections with greater consistency of performance than did those who read the Reading Eye test selections.

III. INTER-RATER RELIABILITY

Inter-rater reliability coefficients based upon scorings of the eye-movement films by the researcher and an assistant are reported in Table XII.

TABLE XII

INTER-RATER RELIABILITY COEFFICIENTS FOR EACH EYE-MOVEMENT MEASURE

N=96

EYE-MOVEMENT MEASURE	r
Number of Fixations	.86
Number of Regressions	.82
Average Span of Recognition	.82
Average Duration of Fixation	.93
Rate with Comprehension	.98

Inspection of Table XII reveals that all of the inter-rater reliability coefficients are substantial in magnitude. The inter-rater reliability coefficient of .98 for rate with comprehension reflects the objectivity of the scoring for that measure. To determine the rate, the scorer has to identify the seven countable lines on the film, mark the beginning and ending points, measure the distance between the marks with a scale provided with the camera, and record the rate indicated by the scale.

The process of counting the number of fixations and number of regressions depends more upon judgments by the scorer, thereby providing more opportunities for variations among scorers. The lower inter-rater reliabilities of .86 for number of fixations and .82 for number of regressions reflect this subjectivity and the greater chance for variation.

The inter-rater reliability coefficient of .82 for average span of recognition reflects the reliability of the fixation score from which it is derived as well as any errors made by the scorers in calculating the average span of recognition scores. The inter-rater reliability coefficient of .93 for average duration of fixation reflects the reliability of the fixation and rate scores from which it is derived as well as any scorer errors.

Subjectivity of Scoring

As suggested in the previous section, the determination of number of fixations and regressions requires a degree of subjective judgment on the part of the scorer. The amount of subjectivity involved is directly related to the quality of the eye-movement film record: the sharper the film image, the more objective the scorer can be. In scoring films made by adults and older children, there is little difficulty. For first graders, however, the abundance of head movements in a typical eye-movement film forces the scorer to judge contantly whether a sharp break in a line represents one fixation interrupted by a head movement or whether it represents two fixations.

The basis for distinguishing between an inter-fixation movement and a head movement is clearly explained in the Reading Eye manual (S. E. Taylor, 1960, pp. 37-39) and is illustrated with several examples. Nevertheless, the difficulty of making such judgments in the numerous borderline cases remained a problem throughout the period of scoring.

In several cases, pupils in this study replaced the typical long return sweep with a series of short right to left inter-fixation movements. The Reading Eye manual (S. E. Taylor, 1960, p. 52) indicates that this pattern is often exhibited by beginning readers, but the manual does not clearly state whether the fixations that follow these right to left movements should be counted as regressions or not counted because they substitute for the return sweep. In such cases, the scorer must make his own decision and any two scorers may decide differently.

Finally, such problems as faulty return sweeps, skipped lines in the reading, rereading of a line, and head movement after finishing a line sometimes made it difficult for the scorers to identify the seven countable lines on a film.

<u>Summary</u>

The inter-rater reliability coefficients computed in this study indicated substantial consistency of agreement between the scorers for each eye-movement measure. Several aspects of the subjectivity of scoring the films were reported.

IV. COMPARISON WITH READING EYE NORMS

Table XIII presents a comparison between the mean eye-movement scores of first grade pupils in this study and mean eye-movement scores reported for first graders in a nationwide norm study with the Reading Eye (S. E. Taylor *et al*., 1960, p. 12). The mean eye-movement scores reported as the Reading Eye norms were obtained in March, 1959, from photographs of 1,028 first grade pupils who were able to read the Reading Eye first grade level test selections with adequate comprehension.

According to Table XIII, the Reading Eye norm for fixations is 224 fixations per one hundred words. Pupils in this study read with a mean of 233 fixations per one hundred words. For regressions, the Reading Eye norm is 52 regressions per one hundred words. Pupils in this study read with a mean of 41 regressions per one hundred words.

TABLE XIII

COMPARISON WITH READING EYE NORMS FOR FIRST GRADE

EYE-MOVEMENT MEASURE	Reading Eye Norms[a] $\bar{X}$ (N=1,028)	Findings of This Study $\bar{X}$ (N=96)
Fixations per 100 Words	224	233[b]
Regressions per 100 Words	52	41[b]
Average Span of Recognition	.45	.46[c]
Average Duration of Fixation	.33	.56[c]
Rate with Comprehension	80	66[b]

[a]From S. E. Taylor et al. (1960, p. 12).

[b]Rounded to nearest whole number.

[c]Rounded to nearest hundredth.

The Reading Eye norm for average span of recognition is .45 words. Pupils in this study had an average span of recognition of .46 words. For average duration of fixation, the Reading Eye norm is .33 seconds per fixation. The pupils read in this study with an average duration of .56 seconds per fixation. Finally, the Reading Eye norm for rate with comprehension is 80 words per minute. The pupils in this study read at a mean rate of 66 words per minute.

Only for average duration of fixation were differences between the Reading Eye norms and the mean eye-movement scores obtained in this study substantial. For number of fixations, number of regressions, average span of recognition, and rate with comprehension, mean scores of the first grade pupils in this study were much like mean scores of the pupils in the nationwide norm study. Fixations of pupils in this study, however, averaged 70 per cent longer in duration than the norm.

While the reason for the much greater average duration of fixation for the pupils in this study than for the pupils in the norm study is unknown, it is suspected that the mean duration scores in this study were overestimated by the extrapolation of a few extreme scores necessary in the transformation of easier selection scores described in Chapter III.

CHAPTER V

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

This chapter summarizes the background, problem, procedures, and findings of this study. In addition, the chapter presents conclusions, implications, and recommendations for further research.

I. SUMMARY

The visual tracking technique was developed to improve reading performance by increasing a learner's perceptual processing efficiency during reading. Evidence of the effectiveness of the technique, however, was limited; and no evidence existed concerning the effectiveness of visual tracking at first grade level where the training might logically have the greatest utility.

This study was designed to provide empirical evidence to support or refute the proposition that visual tracking training will significantly improve first graders' reading performance by improving perceptual processing efficiency. In addition, the study was designed to provide test-retest and inter-rater reliability data for the eye-movement scores of first grade pupils.

Statement of the Problem

The purpose of this study was to determine, for a selected population of first grade children, the effectiveness of visual tracking training as represented by Visual Tracking, a published tracking program. To make this determination, the following problem was investigated: What are the effects of visual tracking training upon five selected components of first graders' silent reading performance measured by eye-movement photography? These five components are: (1) number of fixations, (2) number of regressions, (3) average span of recognition, (4) average duration of fixation, and (5) rate with comprehension.

The following hypotheses were tested in order to answer the primary question posed in this study:

1. There will be no significant differences between the mean performances of the experimental and control groups on the five eye-movement measures.

2. There will be no significant differences between the mean performances of subjects of either sex on any of the five eye-movement measures.

3. There will be no significant interaction between sex and visual tracking training on any of the five

eye-movement measures.

A second purpose of this study was to gather data regarding the test-retest and inter-rater reliabilities of the five eye-movement measures and to gather information regarding adaptive testing procedures when photographing eye movements of first grade children.

Procedures

This study utilized a posttest-only factorial design with two independent variables, treatment and sex. Comparisons were made on the treatment variable between mean scores of experimental and control group subjects; on the sex variable, between mean scores of male and female subjects.

The dependent variable was silent reading performance as represented by each of the following eye-movement measures: (1) number of fixations, (2) number of regressions, (3) average span of recognition, (4) average duration of fixation, and (5) rate with comprehension.

The subjects in this study were all 115 pupils in the five first grade classes at the Weatherbee Elementary School in Hampden, Maine. Sixty-one boys and fifty-four girls were randomly assigned to one of four sections without regard to homeroom membership. Two sections were randomly selected to receive the experimental treatment; the other two sections, to receive the control treatment.

The training period began February 24, 1969, and lasted twelve weeks. Each day, the pupils moved to assigned sections to participate in the control or experimental activity for fifteen minutes. The first grade teachers followed a rotating schedule in supervising the sections so that each teacher spent an equal number of days with each section.

During the fifteen-minute training period each day, pupils in the experimental group used Visual Tracking, a published tracking program. During the same fifteen-minute period each day, pupils in the control group participated in listening activities directed by one of the teachers.

At the conclusion of the twelve-week training period, all pupils were tested in random order with the Reading Eye as they read a fifty-word Reading Eye test selection or one of the fifty-word easier test selections written by the researcher. To collect data for determination of test-retest reliability, retests were administered to forty-four pupils. Films were scored by the researcher and an assistant to determine inter-rater reliability of the scoring.

Findings of the Study

It was found that none of the differences between mean eye-movement scores of the experimental and control groups was statistically significant at the .05 level of confidence. The first hypothesis, therefore, was not rejected.

It was found that differences between mean eye-movement scores of the boys and girls for number of fixations and rate with comprehension were statistically significant at the .05 level. Because the girls in this study read the test selections with significantly fewer fixations and at a significantly faster rate than the boys, the second hypothesis was rejected for number of fixations and rate with comprehension.

For number of regressions, average span of recognition, and average duration of fixation, however, differences between mean scores of the boys and girls were not statistically significant at the .05 level. The second hypothesis, therefore, was not rejected for number of regressions, average span of recognition, and average duration of fixation.

No significant interaction was found between sex and visual tracking training on any of the five eye-movement measures. The third hypothesis, therefore, was not rejected.

Test-retest coefficients for forty-four pupils ranging from .66 to .88 were similar to reliability coefficients reported in earlier studies. The test-retest reliability coefficients in this study were higher than most coefficients reported in previous studies using elementary school pupils.

Inter-rater reliability coefficients ranged from .82 to .98. Since no previous study has reported inter-rater reliability data for eye-movement photography, comparisons with other studies were not possible.

II. CONCLUSIONS

The following conclusions have been drawn regarding the findings of this study:

1. Twelve weeks of visual tracking training with the Visual Tracking program did not result in significantly better reading performance, as measured by eye-movement photography, for the experimental than for the control group. The effectiveness of visual tracking training as a means of improving first graders' reading performance by improving perceptual processing efficiency remains unsupported by empirical evidence.

2. The successful use of eye-movement photography as a criterion measure of reading performance with first

grade pupils has been demonstrated in this study. Equipment and procedure modifications developed in this study contributed to this success.

3. The eye-movement photography technique with the fifty-word test selections read before the Reading Eye has adequate reliability for group measurement of first grade pupils.

4. Inter-rater reliability coefficients were substantial, indicating that the scoring of first graders' eye-movement films is sufficiently objective for group measurement.

5. In general, the mean eye-movement scores of pupils in this study and of pupils in an earlier nation-wide norm study are very similar. Only in average duration of fixation is there a substantial difference between the findings of this study and of the earlier norm study.

III. IMPLICATIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

Effects of Visual Tracking Training

In Chapter II it was proposed that the effectiveness of visual tracking training may depend upon the extent to which cue reduction is a function of familiarity of letter forms or a function of some form of graphemic-phonemic recoding.

The question raised in Chapter II regarding the means by which cue reduction is accomplished was not directly investigated in this study and, therefore, cannot be answered with certainty. Nevertheless, some relationships can be explored on the basis of the findings of this study.

In this study, training with the Visual Tracking program did not result in significantly better reading performance for the experimental group than for the control group. It was suggested in Chapter II that the increased familiarity with letter forms as a result of practice in rapid discrimination of letters in the Visual Tracking program may facilitate the reduction of cues in word recognition. Since there is no evidence that cue reduction occurred to a degree sufficient to produce significant differences in reading performance, the proposition that cue reduction in reading is a function of increased familiarity of letters was not supported in this study.

It was also proposed in Chapter II that if cue reduction is primarily a function of recoding into higher-order units, the visual tracking training may facilitate perceptual processing only for those pupils who are unable to develop such units because of a more basic problem, namely, confusion among letter forms.

A possible indication that the Visual Tracking program may be more effective with some pupils than others is provided by comments of the first grade teachers at the conclusion of the study. In considering these comments, it is recognized that they represent the teachers' subjective reflections rather than conclusions based upon empirical evidence.

While the teachers reported that all pupils in the experimental group improved in their ability to complete the Visual Tracking exercises quickly and accurately, they reported that the visual tracking training seemed especially helpful for the pupils who lacked proficiency in letter recognition and in oculomotor control.

For example, one teacher wrote:

> I feel that Visual Tracking was beneficial to all students but not to all in the same way. . . . With the more immature children, those who did not know their alphabet, it helped them to learn their letters as well as their natural order, left to rightness, and returning to the very next line.

Another teacher stated:

> . . . It [visual tracking training] also helped the children to drop one line at a time and return to the left much faster. I noticed that D--- and J--- were much quicker about that than either D---- or A--- who were in the listening group. The last two were much quicker in recalling events read to them than were the others in their reading group. The four mentioned were in my lowest reading group.

> . . . On the whole I think the program was a big help to the children who did not recognize their letters and had trouble in returning to the left on the following line.

Another teacher wrote:

> I had one or two in this class that needed lots of help with the alphabet and I found that they did recognize their letters much better and faster and knew their placement without seeing the alphabet in front of them.

Examination of the completed Visual Tracking programs and progress charts supported the teachers' comments that all pupils improved in their ability to complete the exercises accurately and quickly. The absence of a significant difference between the experimental and control groups on any of the five components of reading performance measured by eye-movement photography indicates that improvements in tracking ability did not have any measurable effect upon the reading process for most pupils.

The fact that improvement in tracking skill had no measurable effect upon the reading performance of most pupils may indicate that cue reduction is primarily a function of recoding into higher-order units and that the majority of the first grade pupils in this study had already achieved a sufficient level of proficiency in letter recognition to allow the development of such units.

Perhaps the lack of effect upon reading performance also resulted because most of the pupils had already developed adequate left to right movement, ability to stay on the line, and ability to make accurate return sweeps to the next line. Although these findings do not disprove the possible need for oculomotor training in the beginning stages of reading instruction, they do suggest that such training may be unnecessary for the majority of pupils by mid-year of first grade.

Further research is needed to determine the effectiveness of the Visual Tracking program in improving the reading performance of given first grade pupils with specific characteristics, e.g., pupils who read very slowly, pupils who exhibit an unusual number of letter and word reversals, pupils who experience unusual difficulty in learning to recognize letters or words, and pupils who exhibit oculomotor problems.

It was suggested in this chapter that the pupils in this study may have already acquired adequate skill in letter recognition and oculomotor performance and that they were, therefore, already proficient in those skills which the Visual Tracking program purports to develop. Further research is needed to determine the effectiveness of the Visual Tracking program with younger pupils. Perhaps the tracking training could be used effectively as part of an emphasis upon early learning of letter names in a kindergarten or beginning first grade reading program.

Further research is also needed to determine the effectiveness of other types of visual tracking exercises designed to prevent or remedy specific problems in reading. For example, tracking exercises have been developed and used in a clinical setting to provide practice with a letter or short word frequently reversed by a pupil when reading. Such exercises using letters and short words known to be frequently reversed by pupils might be effectively used as a preventative technique with first grade pupils who begin to exhibit a tendency to make many reversals.

Eye-Movement Photography as a Criterion Measure

An indication that eye-movement photography served in this study as an appropriate criterion measure of reading performance is provided by the finding of significant differences on the sex variable. Most studies of reading have found that, on the average, girls read better than boys at first grade level. Two reports from this study support these typical findings regarding sex differences in reading performance at first grade.

First was the finding that girls read the selections with significantly fewer fixations and at a significantly faster rate maintaining adequate comprehension. Second, of the twelve pupils eliminated from the study because of inability to read a test selection with adequate comprehension, nine were boys and three were girls. Because three times as many boys as girls were eliminated from the study for lack of test scores, it is probable that the differences between eye-movement scores for each sex represent minimums. If eye-movement scores could have been obtained for the twelve poor readers who were excluded, it is probable that the differences between means on the sex variable would have been greater. If so, it is also possible that the differences between means of boys and girls for number of regressions and average duration of fixation would then have become statistically significant.

The fact that eye-movement measures of reading performance in this study revealed sex differences in reading typically reported for first graders provides support for the appropriateness of eye-movement photography in measuring the reading performance of groups of first grade pupils.

An implication of these findings is that eye-movement photography could find greater use for groups of first grade pupils in research situations where, because of feasibility and the need for an objective measure of reading performance, the technique is judged appropriate.

It is recommended that researchers using the Reading Eye for group studies consider only the three eye-movement measures that are read directly from the film, namely, number of fixations, number of regressions, and rate with comprehension. The computation of average span of recognition and average duration of fixation may have some value in diagnosis of individual subjects. Since, however, the average span of recognition and average duration of fixation scores are derived from the number of fixations and rate scores, the span and duration scores cannot exceed those scores in reliability or vary independently of them. As a result, comparisons of group means for average span of recognition and average duration of fixation have little meaning.

It is also recommended that researchers using the Reading Eye with young children consider adopting the equipment and procedure modifications developed in this study and that they explore further opportunities to improve the photographic techniques.

Test-Retest Reliability

The test-retest reliability coefficients based upon testing of forty-four pupils indicate that eye-movement photography has adequate reliability for group measurement of first grade pupils. When, however, the reliability coefficients were computed separately for scores of pupils who read the Reading Eye test selections and for the scores of those who read the easier selections, wide differences were found between the coefficients of each group for number of fixations and number of regressions. According to Table XI, page 120, low reliability coefficients of .51 for fixations and .36 for regressions were found for the Reading Eye selections and very high reliability coefficients of .91 for fixations and .90 for regressions were found for the easier selections.

Although the reason for these wide differences in reliability coefficients for the two selections is unknown, several explanations are possible. The wide differences may reflect differences between the groups that read each selection. The group that read the Reading Eye test selections included the better readers in the first grade population, that is, those pupils who were able to perform adequately on the oral pretest. The group that read the easier selections included the poorer readers in the first grade population, that is, those pupils who were unable to read the Reading Eye test selection adequately on the oral pretest.

The better readers performed with greater variability in terms of fixations and regressions than the poorer readers. This greater variability of the better readers may represent the same greater flexibility of eye-movement behavior exhibited by the good readers in Anderson's (1937) study reported in Chapter II.

It is also possible that composition by sex of the groups that read each selection contributed in some way to the wide differences in reliability coefficients. On the basis of performance on the oral pretest, seventeen girls and four boys read the Reading Eye selections before the camera while eleven girls and twelve boys read the easier selections before the camera.

Another possible contributing factor is that 74 per cent of the pupils who read the easier selections read both at one sitting while only 62 per cent of those who read the Reading Eye selections read both at one sitting. Although separate reliability coefficients were not computed for one sitting versus two sittings because of the small number of cases who read at two sittings, it is possible that consistency of eye-movement performance would be higher for two readings at one sitting than for two different sittings.

Finally, the difference in reliability coefficients for the two sets of test selections may have resulted largely from effects of the particular distribution of scores for each of the groups and the small number of subjects in each group. Further investigation might reveal which factor or combination of factors is responsible for the wide differences between the reliability coefficients of the Reading Eye and the easier selections for number of fixations and number of regressions.

Further research is also needed to determine the effects of various testing and scoring procedures and equipment modifications upon young children's eye-movement scores. Such research could serve to further decrease the number of unscorable and marginal films obtained and to improve the test-retest reliability of the technique and inter-rater reliability of the scoring. Also, replication of this study with other populations of first grade pupils would be appropriate.

REFERENCES

Anderson, I. H. Studies in the eye-movements of good and poor readers. Psychological Monographs, 1937, 48, 1-35.

Anderson, I. H., & Dearborn, W. F. The psychology of teaching reading. New York: Ronald Press, 1952.

Ballantine, F. A. Age changes in measures of eye movements in silent reading. In Studies in the psychology of reading. University of Michigan Monographs in Education, No. 4. Ann Arbor: University of Michigan Press, 1951. Pp. 65-111.

Bannatyne, A. D. The transfer from modality perceptual to modality conceptual. In H. K. Smith (Ed.), Perception and reading. Proceedings of the 12th annual convention, International Reading Association, Vol. 12, Pt. 4, Newark, Delaware, 1968. Pp. 7-16.

Bergan, J. R. Perception and reading. University of Kansas Bulletin of Education, 1965, 19, 90-98.

Broom, M. E. The reliability of the reading graph yielded by the Ophthalmograph. School and Society, 1940, 52, 205-208.

Buros, O. K. (Ed.) The sixth mental measurements yearbook. Highland Park, N. J.: Gryphon Press, 1965.

Buswell, G. T. Fundamental reading habits: A study of their development. Supplementary Educational Monographs, 1922, No. 21.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 171-246.

Crawford, J. E. Update of 1963 economic base study--Hampden, Maine. Unpublished manuscript, 1969.

Dember, W. N. Psychology of perception. New York: Holt, Rinehart, & Winston, 1964.

Eurich, A. C. The reliability and validity of photographic eye-movement records. Journal of Educational Psychology, 1933, 24, 118-122. (a)

Eurich, A. C. Additional data on the reliability and validity of photographic eye-movement records. Journal of Educational Psychology, 1933, 24, 380-384. (b)

Fisher, R. A., & Yates, F. Statistical tables for biological, agricultural, and medical research. London: Oliver & Boyd, 1963.

Flanagan, J. C. Units, scores, and norms. In E. F. Lindquist (Ed.), Educational measurement. Washington, D. C.: American Council on Education, 1951. Pp. 695-763.

Frandsen, A. An eye-movement study of objective examination questions. Genetic Psychology Monographs, 1934, 16, 80-136.

Geake, R. R., & Smith, D. E. P. Visual tracking. Ann Arbor, Mich.: Ann Arbor Publishers, 1962. (a)

Geake, R. R., & Smith, D. E. P. Visual tracking manual. Ann Arbor, Mich.: Ann Arbor Publishers, 1962. (b)

Geake, R. R., & Smith, D. E. P. A learning program in perceptual skills. The inter-institutional seminar in child development, collected papers, 1963. Dearborn, Mich.: Edison Institute, 1964, 61-67.

Geyer, J. J. Perceptual systems in reading: The prediction of a temporal eye-voice span. In H. K. Smith (Ed.), Perception and reading. Proceedings of the 12th annual convention, International Reading Association, Vol. 12, Pt. 4, Newark, Delaware, 1968. Pp. 44-53.

Geyer, J. J. Models of perceptual processing in reading. Paper presented at the 14th annual convention of the International Reading Association, Kansas City, Missouri, May, 1969.

Gibson, E. J. Perceptual development. In H. W. Stevenson (Ed.), Child psychology. The Sixty-Second Yearbook of the National Society for the Study of Education, Pt. 1, Chicago: University of Chicago Press, 1963. Pp. 144-195.

Gibson, E. J., Pick, A., Osser, H., & Hammond, M. The role of grapheme-phoneme correspondence in the perception of words. American Journal of Psychology, 1962, 75, 554-570.

Gilbert, L. C. Functional motor efficiency of the eyes and its relationship to reading. University of California Publications in Education, Vol. 11, No. 3, Berkeley: University of California Press, 1953.

Gilbert, L. C. Speed of processing visual stimuli and its relation to reading. Journal of Educational Psychology, 1959, 50, 8-14.

Gilbert, L. C., & Gilbert, D. W. Reading before the eye-movement camera versus reading away from it. Elementary School Journal, 1942, 42, 443-444, 447.

Harris, T. L., Creekmore, M., & Greenman, M. Dot and Jim. Primer reader. Phonetic keys to reading. Oklahoma City: The Economy Company, 1964.

Imus, H. A., Rothney, J. W. M., & Bear, R. M. An evaluation of visual factors in reading. Hanover, N. H.: Dartmouth College Publications, 1938.

Kerlinger, F. N. Foundations of behavioral research. New York: Holt, Rinehart, & Winston, 1964.

Litterer, O. F. An experimental analysis of reading performance. Journal of Experimental Education, 1932, 1, 28-33.

MacIver, C. J., & Geake, R. R. The relationship of reading achievement to visual tracking progress in a fourth grade class. The inter-institutional seminar in child development, collected papers, 1964. Dearborn, Mich.: Edison Institute, 1965, 90-96.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

Morse, W. C. A comparison of the eye-movements of average fifth and seventh grade pupils reading materials of corresponding difficulty. In Studies in the psychology of reading. University of Michigan Monographs in Education, No. 4. Ann Arbor, Mich.: University of Michigan Press, 1951. Pp. 1-64.

Neisser, U. Cognitive psychology. New York: Appleton-Century-Crofts, 1967.

Robinson, H. M., Monroe, M., & Artley, A. S. Fun with our friends. Primer reader. The new basic readers. Chicago: Scott, Foresman, 1962.

Russell, D. H., & Ousley, O. The little white house. Primer reader. The Ginn basic readers. (Rev. ed.) Boston: Ginn, 1961.

Seibert, E. W. Reading reactions for varied types of subject matter. Journal of Experimental Education, 1943, 12, 37-44.

Smith, H. P., & Dechant, E. V. Psychology in teaching reading. Englewood Cliffs, N. J.: Prentice-Hall, 1961.

Spache, G. D. Evaluation of eye-movement photography in reading diagnosis and reading training. In Research and evaluation in college reading. The Ninth Yearbook of the National Reading Conference for Colleges and Adults. Fort Worth: Texas Christian University Press, 1960. Pp. 98-106.

Spache, G. D. Good reading for poor readers. (5th ed.) Champaign, Ill.: Garrard Press, 1966.

Taylor, E. A. Controlled reading. Chicago: University of Chicago Press, 1937.

Taylor, E. A. The fundamental reading skill. Springfield, Ill.: Charles C. Thomas, 1959.

Taylor, S. E. Eye-movement photography with the Reading Eye. (2nd ed.) Huntington, N. Y.: Educational Developmental Laboratories, 1960.

Taylor, S. E. Eye movements in reading: Facts and fallacies. American Educational Research Journal, 1965, 2, 187-202.

Taylor, S. E., Frackenpohl, H., & Pettee, J. L. The validity of eye-movement photography as a measurement of reading performance. EDL Research and Information Bulletin No. 2, Huntington, N. Y.: Educational Developmental Laboratories, 1959.

Taylor, S. E., Frackenpohl, H., & Pettee, J. L. Grade level norms for the components of the fundamental reading skill. EDL Research and Information Bulletin No. 3, Huntington, N. Y.: Educational Developmental Laboratories, 1960.

Tinker, M. A. Eye movements in reading. Journal of Educational Research, 1936, 30, 241-277. (a)

Tinker, M. A. Reliability and validity of eye-movement measures of reading. Journal of Experimental Psychology, 1936, 19, 732-746. (b)

Tinker, M. A. Motor efficiency of the eye as a factor in reading. Journal of Educational Psychology, 1938, 29, 167-174.

Tinker, M. A. The study of eye movements in reading. Psychological Bulletin, 1946, 43, 93-120.

Tinker, M. A. Recent studies of eye movements in reading. Psychological Bulletin, 1958, 55, 215-231.

Tinker, M. A. Bases for effective reading. Minneapolis: University of Minnesota Press, 1965.

Vernon, M. D. Backwardness in reading. Cambridge: Cambridge University Press, 1957.

Vernon, M. D. The perceptual process in reading. Reading Teacher, 1959, 13, 3-8.

Vernon, M. D. The psychology of perception. Baltimore: Penguin Books, 1962.

Vernon, M. D. Approaches to word perception. Education, 1966, 86, 459-463.

Walker, R. The eye movements of good readers. Psychological Monographs, 1933, 44, 95-117.

Wepman, J. M. The modality concept--Including a statement of the perceptual and conceptual levels of learning. In H. K. Smith (Ed.), Perception and reading. Proceedings of the 12th annual convention, International Reading Association, Vol. 12, Pt. 4, Newark, Delaware, 1968. Pp. 1-6.

Woodworth, R. S., & Schlosberg, H. Experimental psychology. New York: Henry Holt, 1954.

APPENDIX A

Sample Exercises From

Visual Tracking*

*Reprinted from Visual Tracking by R. R. Geake & D. E. P. Smith by permission of Donald E. P. Smith. Copyrighted 1962. Permission to reproduce this copyrighted material has been granted by Donald E. P. Smith to the Educational Resources Information Center (ERIC) and to the organization operating under contract with the Office of Education to reproduce ERIC documents. Reproduction by users of any copyrighted material contained in documents disseminated through the ERIC system requires permission of the copyright owner.

Fgn acog zbof hdmi nush pow 1
tryp kif nurc shd. Mo jins pem ruk.

Ingril norftn ratole kihbrtn. Lntrp pifgle zmo lirgt. 2
Incrddl, oftrnecity oplo iglit. Nvloppo rstrintr. Ixplnitri,
Loop aprafrnso brtnr introploppo. Abagnsro sop ippi ojko.

Jeggker, ixqus ufrag. Decte glik egjf beghi ikki abbi. 3
Monognex Glap eppaw. Kasn jes kn b mch fn tleux impeyvent.
Repor grap ippi degho. Gimprwr ixcap mork, improynarzpr.

Peant derna nyws torrnib jerst wuttals plagter. Murs bis 4
prid tolik rige mecol, filt rsd turosp. Ness kgzh wuvvop irp
ponur zoss paftor pox kotall voxabbntuv quog spon kuab lortn.

Seregatti kerp ippeliner iblopper. Oklipper ipswit igglernol.
Incrddl, oftr necity oplo iglit. Nvloppo rstrintr. Ixplnitri. 5
Sop Jifoo aboro. Iglooat nor ippi ros sop iggnorti rstrntr.

Ifligger mort newqapto iblix morsn tspt flopi morcner. 6
Iglox ponur hloffdon strik hloffe ris zmo loox pliggirs.
Ragbaf mort bagc trakt. Mort klappi horatis ignori.

Imponxat ert ftk encrobixer. Opmono rnc lft klixten. 7
Jnopper tsn kgi tvz sji intrdixer. Mort frig, impi optor
Trab ippi jloktar. Oplgnsto kla cakh abbo trst trap rista.

Triak rstv nrm ntyp ecrrbx rsq kljju ns opfl wntve nze. 8
Loi nrtn sprs kcq een ort if sprt ordn slj gqi prse qnv tw opnks.
Ponur zoss paftor ponkot allvo xabbntu vqu ogpon puan lortn.

Painip dre svt eer nmok opr linf expt lkrrp ghr jqs tixibic. 9
Sor mnyxnt pinriku, trlk dfu ntlkoi jtr. Min eopf rqtw.
Tobaj imp inx ort rlat ximgrtr. Trspta clinx pahtr, ipo nrx.

APPENDIX B

Sample Lesson Plans for Control Group

Listening Activities

10. Repeat the Code   Dates Used ______________________

Purpose: To listen attentively in order to be able to accurately reproduce a tapped code.

Materials: None

Directions:

The teacher begins the game by tapping a sound pattern with some object on the side of her desk or other surface. (The teacher's hand should be out of sight so that the pupils cannot determine the pattern visually.) Each child who thinks he can reproduce the pattern exactly raises his hand. When a child is called on by the teacher he tries to reproduce the pattern by tapping on his desk or other surface. The class should be asked to judge whether or not the child reproduced the pattern exactly. If he did, he can be allowed to give the next stimulus pattern. If he did not, another child is chosen to imitate the teacher's pattern.

Patterns should be simple at first and increase in complexity as pupils increase skill in reproducing them. Sound patterns can be made more complex by increasing the number of taps and by adding additional pauses between taps.

Examples: Easy - tap-tap-tap or tap-pause-tap
Hard - tap-pause-tap-tap-tap or tap-tap-pause-tap-pause-tap

11. Prepositional Directions   Dates Used ______________________

Purpose: To listen attentively and follow directions which vary the preposition used.

Materials: Various common classroom items readily available.

Directions:

The teacher or a pupil gives directions to a single child, a small group, or the entire class. The pupils respond to the directions given.

This game is similar to number six but emphasizes attention to the prepositions used rather than following directions in sequence (although sequence can be added to this game, too).

Prepositions such as: on, in, under, over, into, inside, through, by, near, beside, below, off, above, etc. should be used in various ways with different objects and body parts to give simple and complex directions.

Examples: Simple - "Put one hand behind your back and the other hand on your head."
Difficult - "Slide your pencil through the space under your chair and then put it inside the desk of the of the person who sits near the door."

APPENDIX C

Sample Progress Chart from the

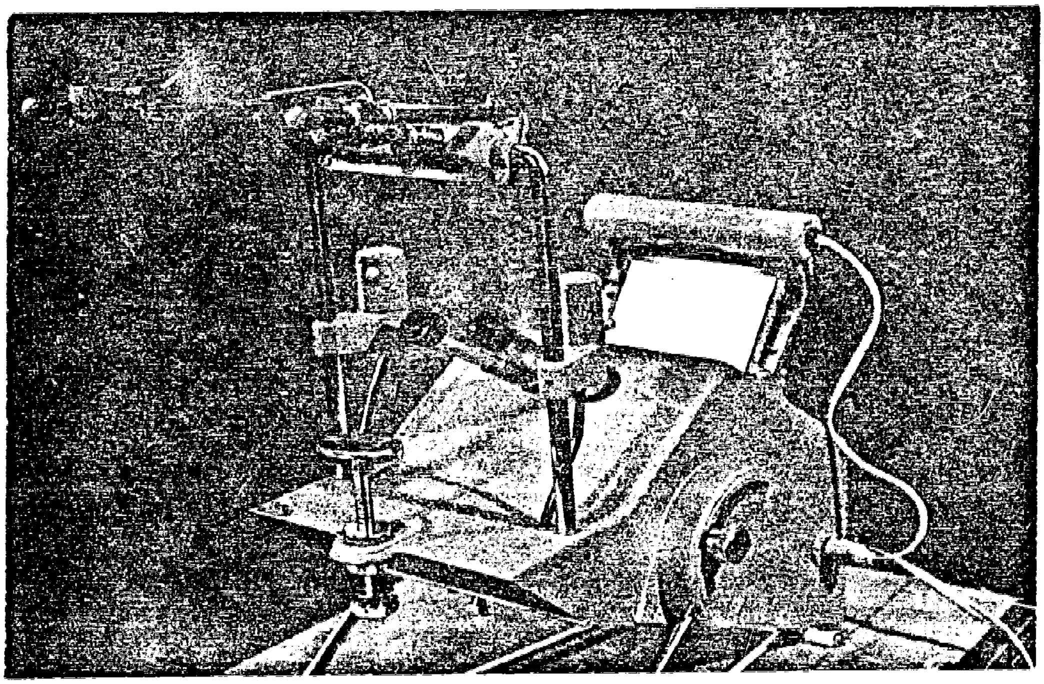
Visual Tracking Program*

20
30
40
50
00
10
20
30
40
50
00
10
20
30
40
50
00
10
20
30
40
50
00

27 28 29 30 31 32 33 34 35 36 37 38 PAGE

APPENDIX D

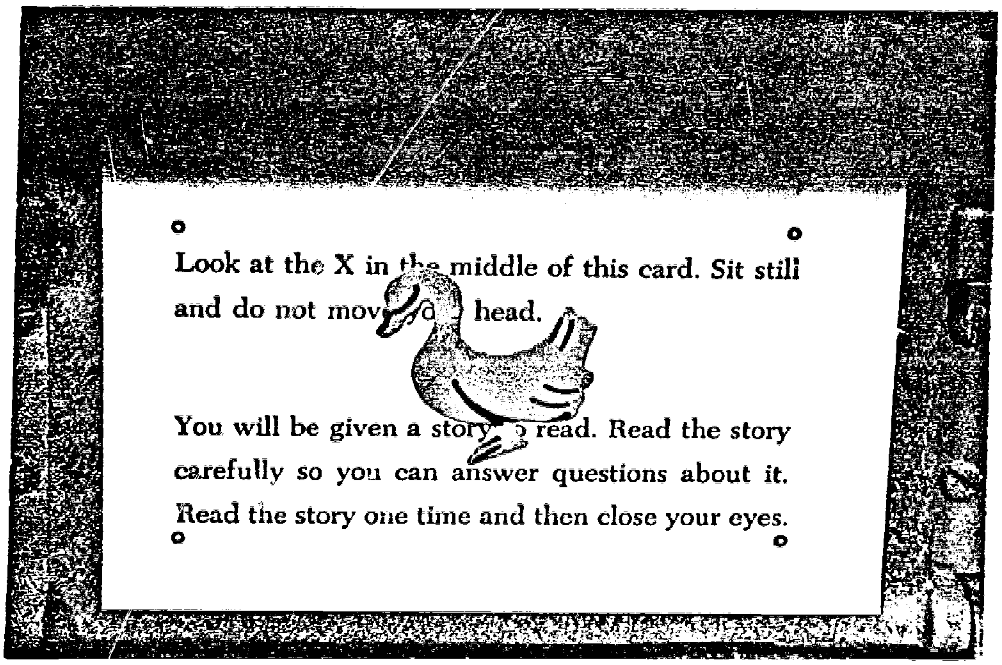
The Reading Eye



APPENDIX E

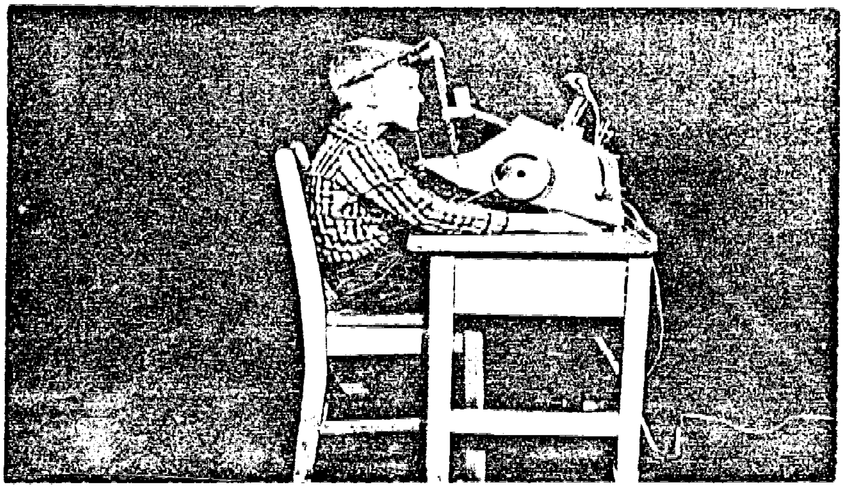
Reading Eye Target Card Showing

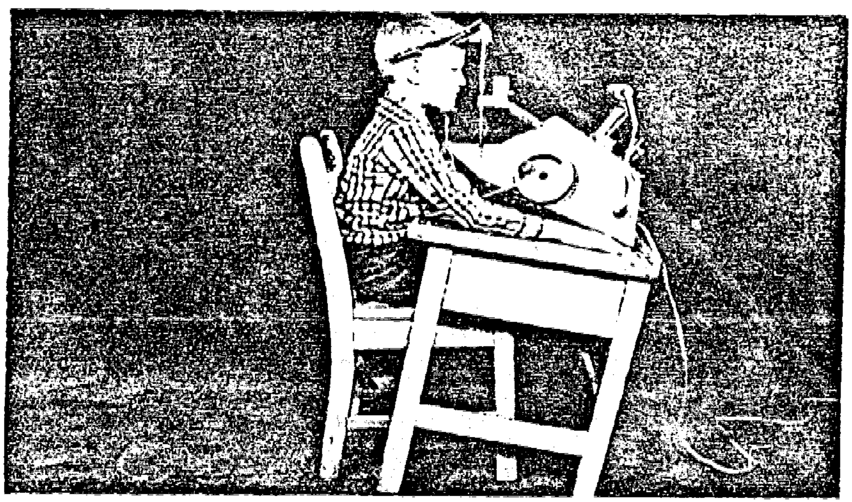
Animal Adaptation


Look at the X in the middle of this card. Sit still
and do not mov head.
You will be given a story read. Read the story
carefully so you can answer questions about it.
Read the story one time and then close your eyes.

APPENDIX F

A Pupil Using the Reading Eye
at a Flat Table and at
the Modified Table





APPENDIX G

Easier Test Selections

and Questions

STORY 1

Jim walked to the dog house.
He saw a little dog in it.
The little dog was white and black.
Jim said, "You are not my dog."
He ran fast to get his mother.
She came to see the new dog.
It was not there. Where did it go?
They looked and looked for the dog.
But they did not see it.

QUESTIONS

1.	Jim walked to school.	No
2.	He saw a little dog.	Yes
3.	The dog was in Jim's dog house.	Yes
4.	The dog was brown and white.	No
5.	It was Jim's dog.	No
6.	Jim ran to tell his father.	No
7.	His mother came to see the dog.	Yes
8.	The little dog wagged his tail when he saw Mother.	No
9.	Jim helped his mother look for the dog.	Yes
10.	They found the dog in the dog house.	No

STORY 2

Bob ran home to the farm.
He wanted to help his father work.
They had to work on the barn.
Bob's big dog jumped up on him.
He wanted Bob to play with him.
Bob said, "I can not stop now.
I must help Father work on the barn.
Run down to the little hen house.
The chickens will play with you."

QUESTIONS

1.	Bob ran to the store.	No
2.	He wanted to help Father.	Yes
3.	They were going to work on the hen house.	No
4.	Bob's dog was small.	No
5.	His dog jumped up on him.	Yes
6.	He wanted Bob to play.	Yes
7.	Bob played with his dog.	No
8.	Bob said he had to help father.	Yes
9.	He said to run to the woods and play.	No
10.	Bob said, "The chickens will play with you."	Yes

STORY 3

Dot had a white kitten.
The kitten liked to run and jump.
The kitten came in to the house.
Dot saw the kitten get a ball.
"I know what you want," said Dot.
"You want to run and get the ball.
I can not play with you now.
I have to go out with Mother.
We can not have fun now."

QUESTIONS

1.	Dot had a little kitten.	Yes
2.	Her kitten was black and white.	No
3.	The kitten liked to sleep.	No
4.	Her kitten came into the house.	Yes
5.	He got a mouse.	No
6.	Dot saw her kitten get a ball.	Yes
7.	She knew what the kitten wanted to do.	Yes
8.	The kitten wanted to play.	Yes
9.	Dot played with the kitten.	No
10.	She had to go with her father.	No

STORY 4

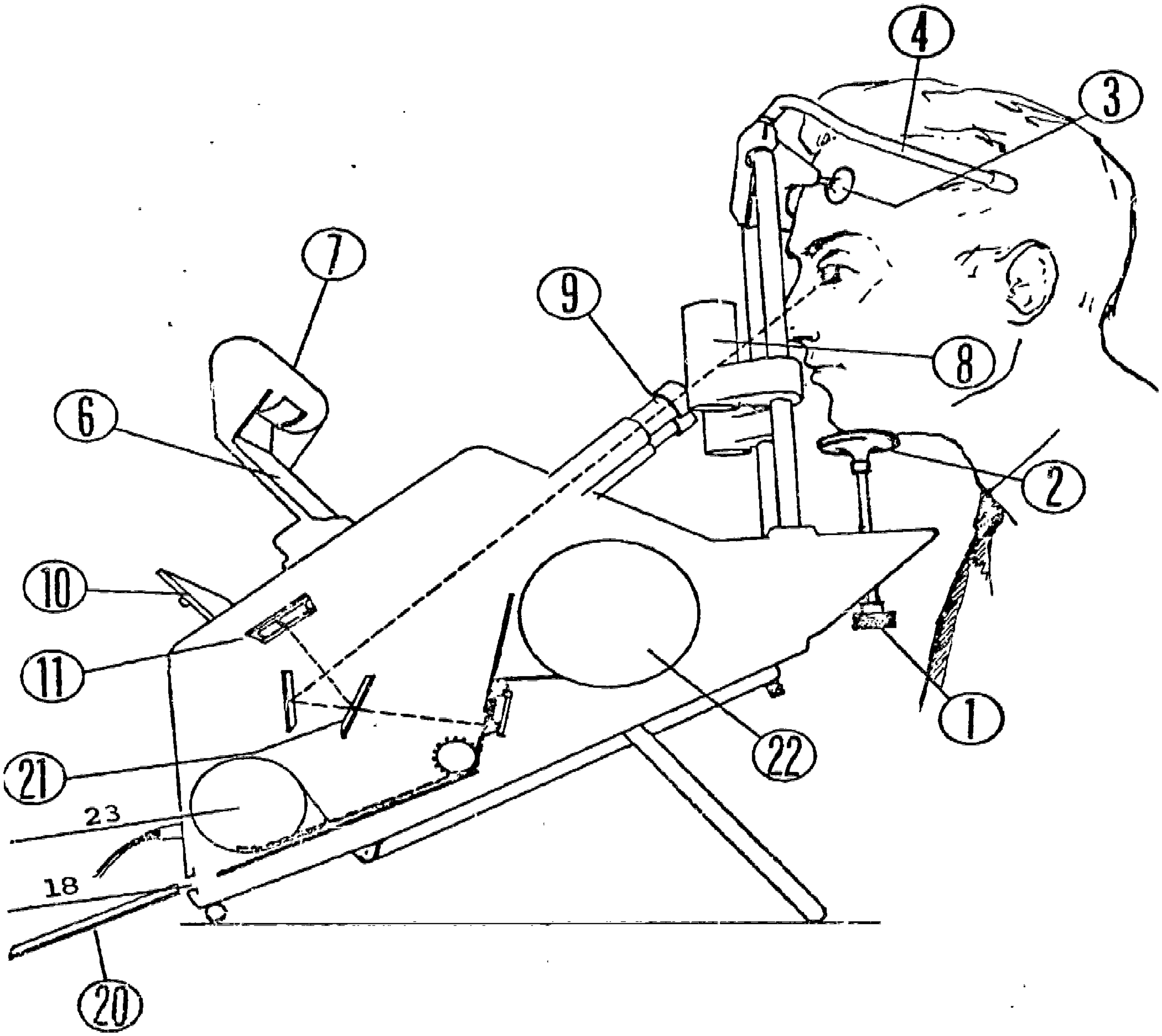
Jane went in the car.
She went with her mother and father.
They went to get a little rabbit.
The rabbit was happy to see Jane.
It hopped to her and jumped up.
Father said, "It is a good rabbit.
Jane and the rabbit are having fun."
They got in the car to go home.
Jane liked her new rabbit.

QUESTIONS

1.	Jane went for a ride.	Yes
2.	She went on a bus.	No
3.	Her mother and father went too.	Yes
4.	They went to get a pony.	No
5.	The rabbit was afraid of Jane.	No
6.	It hid in the cage.	No
7.	Father liked the rabbit.	Yes
8.	Jane and the rabbit had fun.	Yes
9.	They all rode to grandmother's house.	No
10.	Jane liked her new rabbit.	Yes

APPENDIX H

Using the Reading Eye*

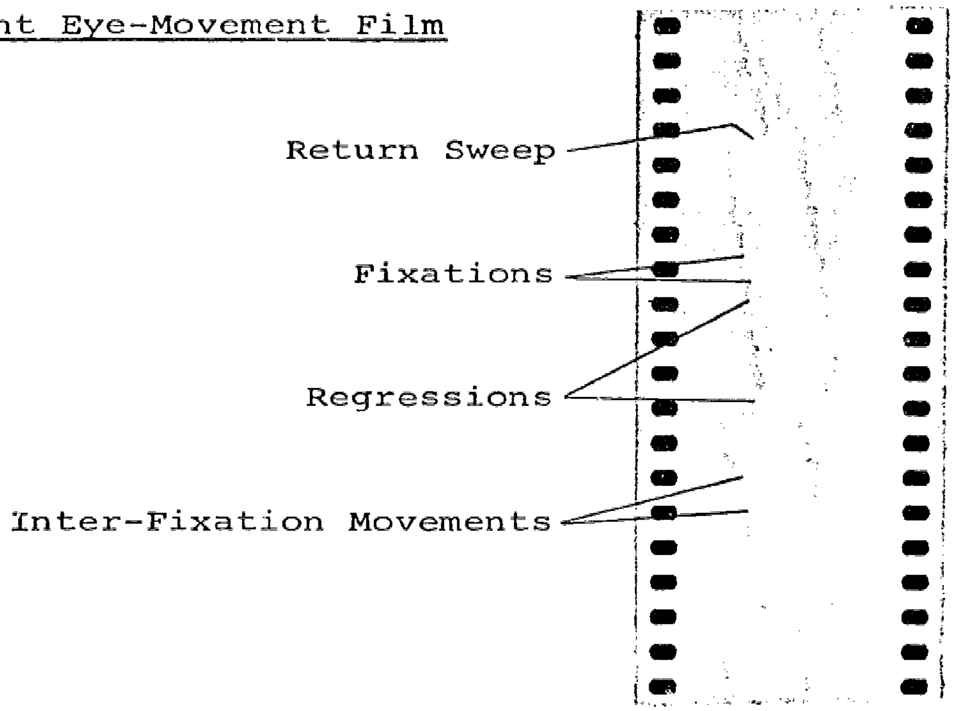


The subject sits, placing his chin on the chinrest (2), his forehead against headrest pads (3). The head steadiers (4) are swung into position against his head. As the subject reads the test selection held in test selection holder (6), illuminated by reading light (7), recording lights (8) cause beads of light to form on the cornea of his eyes. These beads of light pass through telescoping lenses (9) to mirror (21) from which they are deflected upwards to reflex aperture (11) for alignment and focusing or downwards to be recorded on film. To record, reflex aperture door (10) is closed, recording lights dim, and film moves at a continuous rate from film supply spool (22) by recording aperture and into the take-up magazine (23). At the completion of each photograph, the I.D. marker bearing the subject's initials is inserted into the I.D. marker entrance (18), and the initials are flashed onto the graph electronically.

APPENDIX I

Sample Eye-Movement Films

An Excellent Eye-Movement Film



A Poor Eye-Movement Film