

AN INVESTIGATION OF INSTRUCTION TO
READING ACHIEVEMENT MEASUREMENT
FOR NON-ADVANTAGED READING
STUDENTS IN GRADES 5-6

By

MARGARET SHAFFER ATKIN

Bachelor of Arts
Southwestern College
Winfield, Kansas
1963

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1967

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF EDUCATION
July, 1975

Thesis
1975D
A873i
cop. 2

MAY 12 1976

AN INVESTIGATION OF INSTRUCTION TO
READING ACHIEVEMENT MEASUREMENT
FOR NON-ADVANTAGED READING
STUDENTS IN GRADES 5-6

Thesis Approved:

Wm P Ewens

Thesis Adviser

Jm M Sears

Julia L. D. Hahn

Ronald L. Gumble

N N Deuker

Dean of the Graduate College

938874

ACKNOWLEDGMENTS

I wish to thank all of those who have helped bring this study from the abstract to the concrete: Julia McHale, Jim Seals, Ron Gamble, and a special thanks to my major professor W. Price Ewens.

It is with a deep sense of abiding love that I acknowledge the sustaining support that my late husband, J. M. Atkin, gave to me during this study. His wish that I finish it has supplied me with the needed impetus to complete it.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
Significance of the Study	4
Assumptions	5
Statement of the Problem	5
Limitations	5
Definition of Terms	6
Statement of the Hypotheses	6
Organization of the Study	7
II. REVIEW OF LITERATURE	9
III. DESIGN AND METHODOLOGY	22
Subjects	22
Treatment	23
Instrumentation	24
Procedure	25
Statistical Analysis	27
Summary	28
IV. ANALYSIS OF THE DATA	29
Introduction	29
Selection of Groups	29
Collection of Data	32
Analysis of the Data	33
V. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS	40
Summary	40
Conclusions	43
Limitations	44
Recommendations	44
Concluding Statement	45
BIBLIOGRAPHY	46
APPENDIX A	49
APPENDIX B	50

LIST OF TABLES

Table	Page
I. T-Scores of Non-Advantaged Readers	34
II. T-Scores of Advantaged Readers	36
III. T-Scores of Combined Groups	37

CHAPTER I

INTRODUCTION

Since the development of assessment instruments for use in the elementary school setting, it has been assumed that the non-advantaged reader will be able to function in a manner which will allow the test administrator to have an accurate measurement of the student's abilities.

This assumption may arise partially from the fact that directions are standardized and the test administrator is cautioned to strictly adhere to these directions lest the standardization be invalidated. In order to facilitate understanding of standardized tests in education, it would be feasible to define standardized tests. Cronbach (1960) has defined a standardized test as a "systematic procedure for comparing the behavior of two or more persons" (p. 21).

Noll (1965) defines a standardized test as an instrument that has been carefully constructed by experts with specific objectives and purposes. Noll's definition includes the requirements of instructions for administering, scoring, and interpreting so that results may be comparable with norms or averages for different age or grade levels.

The functions of standardized tests are many and varied. One can, in essence, sum up their function by saying that

they should help in decision making. Throughout his school career a student makes many decisions, and many decisions are made about him. The more accurate the information on which a decision is based, the better the decision is likely to be.

Standardized tests provide norms that can be useful in comparing group or individual performance. Tests, per se, do not make decisions, and are not the only information that should be considered in making decisions. In general, standardized tests do provide us with information concerning the probability of the outcome of a decision. In the area of the non-advantaged reading student, the investigation of proper test administration methods is sparse. With the exception of a few individual tests constructed to identify specific disabilities, the non-advantaged reading student's test scores are compared to the norms of the advantaged reading student's. There has been an increased use of standardized tests in education, and for tests to be most helpful in decision making, it is necessary that they be chosen properly, administered correctly, and interpreted accurately.

Currently in the field of reading there are innumerable instruments to use for testing reading in the average classroom. Among those who have constructed these tests there is considerable agreement that the three broad aspects of reading on which information is needed are speed, vocabulary, and comprehension. However, in accepting speed, vocabulary,

and comprehension as the main components toward which measurement is directed, other limitations arise because of the complex nature of the subdivisions. Measurement of rate, reading vocabulary, and knowledge of language used for comprehension may cloud the score. Notwithstanding the current limitations of standardized reading tests, they have a number of positive values for instruction in reading in all schools.

A reading test lends a certain amount of definiteness to thinking about the achievement of a pupil or a group. Comparisons with grade norms, school norms, and national norms may be made.

These kinds of information lend direction to the planning of reading instruction for groups and individuals. Identification of advantaged and non-advantaged readers may be made. The field for the normal and the gifted readers is well supplied with proper tests and testing procedures; the field of the non-advantaged reader remains at the stage of identification.

As each student has individual strengths and weaknesses, it is appropriate to consider the kinds of problems caused by standardized test procedures.

If reading the directions and test items is part of taking the examination, this procedure may be a problem to the non-advantaged reader in several ways. Standardized tests are timed and the non-advantaged reader may not be able to read efficiently with the pressure of time limits.

Incorrect pronunciation of words may cause errors on the test. Visual-perceptual problems may be a cause of a student being classified as non-advantaged, and with this handicap a standardized procedure might not yield a true score.

Because the standardized test procedure may not be altered in any way in order to use the normative data, the child who is a non-advantaged reader may not compare favorably with the norm group.

Significance of the Study

The significance of this study lies in its attempt to find a better method to assist non-advantaged reading students in the process of test taking.

The child who has been inaccurately measured may be handicapped in the classroom in several ways. In comparing his score with established norms, it may appear that he does not have the intellectual capabilities to produce school work within the normal range. He may be demoted, in a sense, to school work of a lower grade. Although the work may be easier for him to perform, the content may be too elementary to hold his interest and motivate him.

There are other facets to be considered. The child may suffer an emotional overlay in the area of his self-concept. Because he has been mismeasured and his problem has not been properly diagnosed, frustration and anger may become a part of his personality make-up.

This study is to determine if different methods of administration of standardized test directions may be desirable when such tests are to be given to non-advantaged readers.

Assumptions

Because of the nature of this study we may assume that:

1. The data are linear in nature, thereby meeting the restrictions of the statistical technique, the T-score formula.
2. The administration of the tests was uniform.

Statement of the Problem

The problem of this study was to determine if identified non-advantaged reading students would obtain a significantly different score when given the Stanford Reading Test by regular procedure as opposed to the directions administered by tape.

When students are classified as non-advantaged reading students a more accurate measurement may be obtained by administering a reading test by other methods than the standardized procedure as given in the test manual.

Limitations

1. The reliability of the measuring instruments will to some degree affect the reliability of any conclusions drawn in this study.

2. Other than the students being identified as non-advantaged readers of the fifth and sixth grade of the Pawhuska Elementary School, Pawhuska, Oklahoma, there is no attempt to control for other student differences.

Definition of Terms

1. The non-advantaged reading student was identified by using Bond and Tinker's (1968) formula, years in school multiplied by I. Q. plus one, for the student's reading expectancy level. The student's expectancy score was then subtracted from his score on the reading section of the Stanford Achievement Test on file at the school. If the resultant score was one and one-half years or more below his expectancy score, he was considered a non-advantaged reader for this study.

Statement of the Hypotheses

This study attempted to determine if fifth and sixth grade students who are non-advantaged readers would have significantly higher scores on standardized reading tests when the teacher used a method of administration other than the one used in the manual. More specifically, will non-advantaged reading students have significantly higher scores if the directions of a reading test are given to them by tape.

The statement of this in the form of the null hypothesis is as follows:

There is no significant difference in the scores of non-advantaged reading students of the fifth and sixth grades when given a standardized reading test using regular directions from the manual as opposed to directions given by tape.

In addition, identified advantaged reading students of the fifth and sixth grades were tested and an attempt made to determine if the advantaged readers would have significantly higher scores on a reading test if the administration of the test was by tape as opposed to regular procedure.

The second hypothesis stated in the null is as follows:

There is no significant difference in the scores of advantaged reading students of the fifth and sixth grades when given a standardized reading test using regular directions from the manual as opposed to directions given by tape.

The hypotheses were tested at the .05 level of significance using the T-score formula.

Organization of the Study

The present chapter includes an introduction to the problem, the significance of the study, a statement of the problem, definition of terms, hypotheses, limitations, and assumptions.

Chapter II contains a review of the research literature pertinent to this study. Chapter III describes the subjects, treatments, instrumentation, and analysis of the data. Chapter IV contains the findings and a discussion of the

results of the study. Chapter V includes conclusions of the study and implication for further research.

CHAPTER II

REVIEW OF THE LITERATURE

The purpose of this chapter will be to review pertinent research literature related to this study. The studies included here reflect the present state of current research as well as earlier studies.

Frederick L. Westover (1958), University of Alabama, felt it would be worthwhile to investigate the differences among students in their performances on tests according to the modes used in administering tests.

Westover compared listening and checking performance with reading and checking performance on 198 students in the College of Education of the University of Alabama in a course in elementary psychology. The students were distributed among seven sections over a period of three semesters and one summer session. Most of the students were sophomores, but there was a considerable number of junior, senior, and graduate special students.

The students used as a textbook Ruch's Psychology and Life, third edition. They used, also, the workbook that accompanies this textbook and were required to fill out the self-tests based on the text. The class periods were spent in informal supplementary discussion by the instructor of the

topics in the textbook, with informal classroom discussion. The self-tests were marked, and any questions raised were discussed.

Following the presentation of each two chapters a test was given covering the chapters. This test was composed of forty objective questions, some true-false and some multiple-choice. Twenty of the questions were in one form, ten on each of the two chapters, and twenty were in the other form. A total of eight listening and reading tests were administered. These questions were drawn at random from a library of test questions which were similar to, but not the same as, the questions in the self-test in the workbook. Judged on the similarity of the scores earned, the two forms of the test were considered to be of approximately equal difficulty.

One form of each test on the two chapters being tested was administered by the instructor reading the questions aloud twice while the students listened and wrote the letters of the answers on an answer sheet. When requested, the instructor would read a question a third time but no more. The other form of the test on each of the two chapters was administered by giving each student a mimeographed copy of the test which the student read silently and then wrote the letters of the answer on the answer sheet.

The forms of the test and methods of administration were alternated, sometimes one form then the other being given first. First, listening preceded reading, and then the reverse. This alternation was intended to rotate out any

influence of position in the order of employing the two kinds of testing, and also, any chance differences in the difficulty of the alternate forms of the tests. At the close of each testing period the students were asked to indicate which kind of test administration they preferred, listening or reading.

In a few instances the questions were administered by means of a tape recorder. The students indicated to Westover that they liked this method of presentation because of its clearness and greater volume.

The mean difference on all tests presented by the two methods was .02 of a point in favor of listening out of a maximum possible difference of 20 points. Students who preferred listening to reading had a mean difference of .001 of a point in favor of listening over a maximum possible difference of 20 points. Students who had earned the highest marks in the course were ranked in scholarship and were compared on the excess of their reading scores over their listening scores. The rank order coefficient of correlation was .17.

Students who had earned the lowest marks in the course were ranked in the order of their poorness in scholarship and then compared with the excess of their listening scores over their reading scores. The rank order coefficient of correlation was -.19.

The conclusions drawn from the investigation were that there were no group differences among college students in

performance on tests administered by listening and by reading, and that listening tests are about equally as fair as reading tests to students of high and low scholarship. However Westover also concluded that some students did show consistent differences in performance on similar tests administered by listening and by reading. Westover (1958) used the following as an example:

The record of a student who had received a failing mark in the course in elementary psychology was examined. It was found that if this student's term average had been calculated on the basis of his scores on the listening tests, instead of both listening and reading tests, he would have received a passing mark. In other words, this student had been awarded a failing mark, not because of his inability to understand psychology as measured by his answers to spoken questions, but by his poor performance in answering questions which he read (p. 44).

Westover's investigation illustrated a fact that there may be consistent and important differences among individuals in such performance. It is also speculative that the reading of the test items permits greater possibility of cheating.

Using the same textbook that Westover used in his investigation, Charles L. Odom and Ray W. Miles (1951) also studied a group of students in a course in General Elementary Psychology. The specific questions in the Odom-Miles study (1951, pp. 470-477) were stated as follows:

1. Can students make higher scores on achievement tests of the true-false type in General Psychology when the questions are read to them by the instructor or when they are presented in mimeographed form?
2. Is there a relationship between the level of a student's achievement in General Psychology and the question of whether he will do better on an

achievement test of the true-false type when they are presented in a mimeographed form?

The subjects used in the Odom-Miles study were 200 sophomore students in the College of Education of Southwestern Louisiana Institute. They ranged in ability from very low to fairly high and were composed of an approximately equal number of men and women. These students represented the student group over a period of three years.

The tests used consisted of 15 sets of true-false questions, ranging in length from 50 to 75 questions per set, depending upon the nature and the length of materials to be covered, each test based upon a chapter or a group of chapters. One of the tests was administered each week of the semester in the course in General Elementary Psychology. In addition, each student took a mid-semester and final examination, each consisting of 100 multiple-choice type questions covering half of the course.

To 100 of the students (oral group) the true-false type tests were administered by the instructors reading aloud to the class each question twice, without comment, and then going on to read the next question until the particular test was completed. The other 100 students (visual group) were permitted to take the same test but used a mimeographed form so they might read them to themselves. The second group of 100 students were matched with the first 100 students by matching their scores with those on the mid-semester and final examinations. These multiple-choice tests were taken by all students in both groups in mimeographed form.

In answer to the two questions proposed by Odom and Miles they concluded that:

1. In general, students in General Psychology do about equally well on achievement tests of the true-false type whether the questions are read to them by the instructor or presented to them in mimeographed form. Whatever difference there may be would seem to be in favor of oral presentation.
2. There is a difference between the performance of superior and inferior students on true-false type questions on achievement tests with regard to the importance of method of presentation. Superior students appear to do equally well whether they have an opportunity to read the questions themselves, with some indication that they score slightly higher when they read the questions themselves. Inferior students, on the other hand, do markedly better when questions are read to them by the instructor (p. 478).

In 1932 in a study by L. B. Knox and V. M. Sims of the University of Alabama concerned with the reliability and validity of orally presented multiple-response tests it was concluded that:

1. Multiple-response tests presented orally are but slightly more difficult than the same tests presented visually.
2. Multiple-response tests may be presented orally without seriously reducing the reliability.
3. Multiple-response tests presented orally tend to measure that which is measured by the same tests presented visually.
4. There is no improvement in the reliability of oral multiple-response tests when they are corrected for guessing, but such correction seems to slightly increase the correlation with the visual test.
5. Four or five responses seem superior to three responses when the tests are presented orally.
6. Until further evidence is presented it seems wise to recommend for oral presentation the five-response test with no correction for guessing (pp. 72-73).

For the Sims-Knox investigation, four forms of the Thorndike Test of Work Knowledge, a one-hundred item five-response test, were administered to a group of one hundred students enrolled in grades nine through 12. The four forms were all administered within a period of one month. Form A, given first, was administered in the usual manner, Thorndike's directions being observed. For Form B, the second test administered, two of the wrong choices given for each test word were eliminated and the test given orally as a three-response test. One of the wrong choices for each test word in Form C was eliminated, and this test, the oral four-response test, was given orally for the third test. Form D, administered last was presented orally with all the choices as the five-response test.

The tests were scored by two methods, first by correcting for guessing by means of the usual formula, and secondly without any correction for guessing, the score being the number of correct responses.

In another investigation, Thomas H. Briggs and George H. Armacost (1933) of Columbia University presented orally an important class topic to a class in junior high school. The presentations were made in two class sessions. At the end of each presentation the students were given an oral true-false test of 50 items. Test A, presented to 49 students after a 90 minute lecture, had a range of scores from 24 to 50. Mean = 43.8; standard deviation = 5.90. Test B, 50 true-false items, presented immediately before the second lecture

had a range of scores from -3 to 37. Mean = 25.54; standard deviation = 7.59. Test C, the same test presented immediately after the lecture, had a range of 16 to 47. Mean = 36.44; standard deviation = 6.63.

The authors concluded, both from the statistical data and from their impressions, that the oral true-false test as a measure of immediate recall compares very favorably with such a test presented in visual form. They also concluded that if similar results are found for recall after longer periods, the labor of printing or mimeographing such tests can safely be abandoned.

Stumpf (1928-1931) in two studies calculated the coefficients of correlation of students' scores on tests administered by listening and by reading with their scores on an intelligence test. He found a somewhat higher relationship ($r = .36$) between scores and mental ability ($r = .25$). He concluded that performance on the listening tests yield a truer index of the students' capacity for achievement.

In an experiment using 39 seventh grade pupils as subjects, Caryl Utigard (1962) of the University of Washington reported no difference between results of tests following stories read by the teacher and on tests following stories pre-recorded on tape by the same teacher. The tests were teacher-made from the material presented.

F. Craig Johnson and Kenneth Frandsen (1963) administered the Brown-Carlsen Listening Test to 2400 college freshmen by tape, live, and on film. The best results were

obtained from the use of tape while the film yielded the least satisfactory results. The authors reported that the lecture portion of the Brown-Carlsen does not appear to measure the same skills as the remainder of the test.

John Alfred Jones (1961), also using the Brown-Carlsen Listening Test, reported "very low correlations" between listening test results and college instructors' ratings of listening ability. Using two groups of 49 college speech students, he found no significant differences between a live and a taped presentation of the Brown-Carlsen Test.

Since the identification of the non-advantaged reader is spread over the wide spectrum of non-readers, the writer feels it is appropriate and necessary to discuss the background of some of the various identifications.

It is fairly common in schools to give reading tests to determine if children are reading at their proper academic level. The teacher may use the test scores as a definition of a reading disability. This may lead to erroneous conclusions as to the abilities of the child, for there may be other factors leading to low achievement scores of the disadvantaged, perhaps methods of administration.

Apparently, reading disabilities have been recognized in the literature for over a century (Crawford, 1966). The perceptually handicapped child, the brain damaged child, the child with delayed development of perceptual-motor skills, an emotionally disturbed child, the hyperactive child--all, at times may be defined as having a reading disability. These

definitions are in addition to the child who is economically and socially disadvantaged.

Currently there is much literature concerning the children who have learning disabilities. Although Crawford does not label children with specific disabilities as being disadvantaged, he describes many types of disabilities and their apparent results in the classroom setting.

Crawford states that "Reading disability has long been the Achilles heel for many boys and girls in school" (p. 23). He quotes from The British Medical Journal, 1869, from an article by W. P. Morgan who had written on a case that he thought to be word-blindness. Crawford also draws attention to a survey done in England by Hinselwood of acquired word blindness in adults. Recognition of such difficulties apparently was rare in that century, perhaps because by lack of accurate neurological insight and knowledge. Bereiter and Engleman (1966) found that severely economically deprived children do not perceive language as being comprised of common elements, separate words which may be combined and transformed. Rose M. Bromwich (1971) in a paper delivered at N.C.T.E. National Convention in Las Vegas stated that many children from inner city and poor rural communities have language but it is not recognized as such in the school setting. She says their language is used effectually with their peers and in a one to one relationship with an attentive adult. . .

Therefore these children's (disadvantaged) major educational problem is not the absence of language itself, but it may be that the school, the teacher, the curriculum does not provide him with the proper

opportunities to use his language for learning and for his intellectual growth (p. 21).

In the book, The Disadvantaged Child, Frost and Hawkes (1970) have headed Part One, "The Disadvantaged: Euphemism for the Poor." Six authors--Norman L. Friedman, Carmichael and Charles V. Hamilton, Richard L. Tobin, Peter Farb, and Peter Schrag--have given their definitions of the disadvantaged. Friedman in his article, "Cultural Deprivation: A Commentary in the Sociology of Knowledge," states that,

'Culturally deprived' was a popular image that was able to rally the support of varied interests in order to produce needed legislation and experimental programs . . . and culturally disadvantaged is a similar label that appears to be displacing 'culturally deprived' in academic usage in 1967 (p. 6).

Friedman's article points out that the field of the "culturally deprived" child is limitless as viewed by different cultures. Carmichael and Hamilton (1967) feel that oppression of the minority groups, particularly the Negroes, is the cause of the disadvantagedness of these groups.

Tobin (1968), in "One Million Migrants" states that "exploitation has been and continues to be one of the saddest and most perplexing of all America's social woes (p. 30).

Peter Farb (1968), in "The American Indian: A Portrait in Limbo," points out that the 550,000 American Indians have lived on approximately 200 reservations as poverty-stricken islands surrounded by an ocean of American bounty.

"Appalachia: Again the Forgotten Land" by Peter Schrag (1968) states throughout the article that, although

Appalachia has medium requirements met for the disadvantaged, exploitation of the disadvantaged is still the primary concern. Schrag (1968) quotes Perley F. Ayer, the chairman of the Council of the Southern Mountain, "People expected this thing to be solved in six months, but in education alone we are 5,000,000 years behind" (p. 48).

There is a theme of agreement within the five above mentioned articles as to the definition of the disadvantaged: poverty and exploitation.

Birch and Gussow (1970) in reference to the educationally disadvantaged child argue that our identification of these students can be made, in most instances, before the child is born. Poor maternal health, poor maternal growth, and poor maternal nutrition appear to start the cycle that ends in the definition of a disadvantaged student.

This idea is partially supported by Horn (1970) who says that the assumption is often made that economic poverty alone is the prime cause of educational disadvantage; in fact, proper diagnosis of reading difficulties reveals a wide range of causes . . . the state of disadvantage is a relative matter and may cover a wide range of situations.

A vast amount of work has been done to describe the impact of visual stimuli on the perceptually disadvantaged child. Marianne Frostig (1972) notes that

if certain perceptual skills are not fully automatized, the child's school learning may be permanently slowed or disabled, as most academic subjects have a hierarchical structure . . . the more automatic and subconscious the child's perception are when he reads, the more fluent he

will be and the more he will be able to direct his attention to the thoughts expressed in sentences (p. 57).

Farnhan-Diggory's (1970) experiments showed that ghetto children frequently have learning problems caused by coexisting deficits in visual perception and in the simultaneous synthesis of stimuli. She also found that the children could be helped by a program which combined training in language and action with training in perception. The combination of such training apparently brought added development to the disadvantaged child.

Studies in recent years reveal nearly a quarter of the children in many communities are not able to read at their proper academic level. These boys and girls are likely to make poor showings on any test demanding ability to read (Frost and Hawkes, 1970).

The previous varied reviews indicate the wealth of knowledge available in the area of the disadvantaged student, but it also indicated a lack of investigation in the specific area of test administration methods for the disadvantaged reader. There was no single item found that dealt with administration of reading tests to non-advantaged readers, per se. This appears to be sufficient evidence to warrant some investigation of this particular area and the writer would hope to take one small part of this problem and make suggestions, supported by empirical data, which will provide additional light on the subject.

CHAPTER III

DESIGN AND METHODOLOGY

As concluded by research of the literature, little has been done in the area of seeking different methods of reading test administration for the non-advantaged reader. The purpose of this study is to determine if there may be better methods, score-wise to administer a reading test to non-advantaged reading students than the method prescribed in the test manual--more specifically, will the identified non-advantaged reading students be able to obtain higher scores when taking a standardized reading test by tape as opposed to regular administration procedure.

Subjects

Non-advantaged students were identified by using Bond and Tinker's (1968) formula, years in school multiplied by I.Q., plus one. This formula determines the student's expectancy level. The Stanford Reading scores from the Stanford Achievement Tests and the Otis I.Q. scores of the fifth and sixth grade students were used. These were on record in the school files from the year 1972. The student's expectancy score was subtracted from his score on the reading section of the Stanford Achievement Tests. If the resultant

score was one and one-half years or more below his reading expectancy score he was considered a non-advantaged reader for this study.

After identification of the non-advantaged students, their names were alphabetically arranged, and they were divided into two equal groups by random sampling using Table XVII (Ten Thousand Randomly Assorted Digits) from Statistics for the Behavioral Sciences by Wyatt and Bridges (1967). After selection of the two groups, proof of random sampling was obtained by using the T-score formula to determine if there were significant differences in the mean of the reading scores.

When the non-advantaged students had been identified, the remaining students with scores one and one-half years or more above his reading expectancy were considered advantaged students. Sixty students were selected from this group, again using the alphabetical arrangement of names and random selection as described above.

Treatment

The four groups were given the Stanford Reading Tests on two successive Wednesdays in April, 1973. Form W was used for the first testing and Form X was used for the second testing.

One group of the non-advantaged and one group of the advantaged was administered the test by tape on the first Wednesday and the other groups of non-advantaged and

advantaged students were given the test by regular classroom procedure. The following week the test administration procedure was reversed in the groups. The groups were not tested together (i.e., the non-advantaged and the advantaged were not together when tested). There were four testing sessions each week, these being supervised by the writer. At no time did the classroom teacher take part in the testing procedure.

The T-score formula (Appendix B) was used to determine the significant difference of test administration of both the non-advantaged and the advantaged students.

The second form of the test, Form X, was administered a week later than the first form, Form W. In order to balance out a possible carry over effect from prior experience, a T-score formula was used to determine if there were any significant differences in the taped procedure and the regular procedure of administration when using the combined scores of the taped groups as opposed to the combined scores of the groups using the regular administration.

The testing was done in a classroom especially arranged by the school personnel. The test was not preannounced to the students.

Instrumentation

The instruments used in this study were reading achievement tests. The Stanford Reading Tests, Intermediate II, forms W and X, were constructed by Truman L. Kelley, Richard

Madden, Eric F. Gardner, and Herbert C. Rudman (1964) to ". . . analyze group differences in the subject and also differences in the abilities of individual pupils" (p. 17).

Intermediate II Reading Tests are primarily designed for use from the middle of Grade 5 to the end of Grade 6. They consist of tests in Word Meaning and Paragraph Meaning.

Word Meaning is a 48 item test in which the subject responds by marking a space whose number corresponds to the correct answer. It is designed to provide an objective measure of the student's knowledge of the definition of words.

Paragraph Meaning is a 64 item test in which the subject responds by marking a space whose number corresponds to the correct answer. It is designed to provide an objective measure of the student's comprehension of ideas presented in groups of words.

Statistical data concerning the reliability of the instrument was computed using the split-half reliability coefficients and were corrected with the Spearman-Brown Prophecy Formula. The table which was taken from the manual of directions also reports the standard error of measurement and an estimate of Kuder-Richardson Formula 20 using Saupe's Formula (Appendix A).

Procedure

The four groups were given the Stanford Reading Tests, Forms W and X, on two successive Wednesdays in May, 1973.

Form W was used for the first testing and Form X was used for the second testing.

One group of the non-advantaged and one group of the advantaged were given the test by regular classroom procedure. The following week the test administration procedure was reversed in the groups. There were four testing sessions each week.

The tapes were prepared by the writer. Two recorders were used in the event that one might fail during the testing. The writer read the instructions onto the recorders and they were then edited for errors. There were no errors, nor did a recorder fail during administration.

The T-score formula was used to determine the significant difference at the .05 level and the .01 level of both the non-advantaged and the advantaged students. This permitted testing the hypotheses to determine if the students would score better by tape administration than by regular classroom administration or vice versa. It should be noted that the second form of the test, Form X, was administered a week later than the first form, Form W.

The first step was the identification of the non-advantaged students of the fifth and sixth grades of the Pawhuska Elementary School System, Pawhuska, Oklahoma, by the previously described method.

After identification of the non-advantaged students, their names were alphabetically arranged and divided into two equal groups by random sampling as previously described.

Proof of random sampling was obtained by using the T-score formula to determine if there were significant differences in the mean reading scores at the .05 level and the .01 level.

The remaining population with scores one and one-half years or more above their reading expectancy level, was considered advantaged readers. Sixty students were selected from this group using the alphabetical arrangement of names and random selection as described above.

Statistical Analysis

The statistical analysis was performed at the Oklahoma State University Computer Center. Three computations using the T-score formula yielded the information for the study.

The first computation yielded the difference in the scores when comparing the regular method of administration of reading tests to disadvantaged reading students, as opposed to tape administration.

The second computation yielded the difference in the scores when comparing the regular method of administration of reading tests to advantaged reading students as opposed to tape administration.

The third computation yielded the difference of the combined scores of the two groups of disadvantaged readers with the combined scores of the two groups of advantaged readers.

Summary

Chapter III has presented a description of the design, methodology, and purpose of the present study.

The subjects were selected fifth and sixth grade students of the Pawhuska Elementary School. These subjects were given the Stanford Reading Achievement Tests, Forms W and X, in April, 1973.

The statistical technique employed was the T-score formula.

CHAPTER IV

ANALYSIS OF THE DATA

Introduction

The purpose of this chapter is to report the procedure of this study, the analysis of the data, and the findings. The procedure will be reported first. Following that will be the analysis of the data. Finally, the results of this study will be discussed.

Selection of Groups

The fifth and sixth grades of the Pawhuska Elementary School, Pawhuska, Oklahoma, in 1973 were selected as the population group from which the sample was drawn. There were 357 students in the two grades, 162 in the fifth grade and 195 in the sixth grade.

The non-advantaged students were identified by using Bond and Tinker's (1968) formula, years in school multiplied by intelligence quotient, plus one. This formula determined each student's reading expectancy level.

The school had on file the students' I.Q. scores taken from the Otis-Lennon Mental Ability Test (1967). These tests were administered in January, 1972, and the consequent I.Q. scores were used in the Bond and Tinker formula.

Also on file at the school were the scores from the Stanford Achievement Tests (1964). The tests were administered in September, 1972, and the scores from the Reading section of the test were used with the expectancy score to make the classification of non-advantaged or advantaged. Each student's expectancy score was subtracted from his reading score. If the resultant score was one and one-half years or more below his reading expectancy score he was considered a non-advantaged reader for this study. Any students with a reading expectancy score one and one-half years above his expectancy score was considered an advantaged reader for this study. The one and one-half year level above and below the expectancy level was considered a reasonable cut off point for the scores, thus ruling out scores that might be borderline.

After identification of 60 non-advantaged students, their names were alphabetically arranged and numbered consecutively from 01 to 60. They were then divided into two equal groups of 30 by random sampling using Table XVII (Ten Thousand Randomly Assorted Digits) from Statistics for the Sciences by Wyatt and Bridges (1967).

The digits in Table XVII have been entered in random fashion. To make a random selection, a point of reference was established by pointing to one of the groups and reading the first four digits of the group selected. The first two digits were used to determine the row and the second two digits to determine the column for selecting the first member of the sample. The first group selected was 40914, located

at row 25, column 80-84. Using the first four digits of 40914, row 40 column 91 was selected and the five digit numbers were recorded to the end of the row. At the end of the row, the selections were continued with row 41, column 1.

Since the measures of the group from which the sample was drawn were two-digit numbers, the random digits were grouped into sets of two-digits each. Each set of digits having any value less than 60 was underlined until there were 30 different underlined sets. These sets were arranged in numerical order and then assigned to the corresponding number of the alphabetical arrangement of names. This group was used as the first group of the non-advantaged readers. The remaining 30 in the non-advantaged group were used as the second non-advantaged group.

Sixty students were selected from the advantaged group, again using the random selection arrangement as described above.

After selection of the two groups of the non-advantaged readers and the two groups of the advantaged readers, proof of random sampling was obtained by using the T-score formula to determine if there was a significant difference of the mean of the scores between the two groups of the non-advantaged reading students. This method of proof of random sampling was also used with the advantaged reading students. There was no significant difference in the scores of the two groups of the non-advantaged reading students or between the scores of the two groups of advantaged students.

Tapes were prepared by the writer for administration of the Stanford Reading Tests. Two recorders were used in the event that one might fail during the testing. Instructions for administration of test Form W were read onto the recorders and they were edited for errors. There were no errors, and the recorder did not fail during the administration. The same tape was used for administration of Form X, as instructions were identical.

Collection of Data

The testing was done in a classroom especially arranged by the school personnel. The test was not preannounced to the students.

The four groups were given the Stanford Reading Tests on two successive Wednesdays in April. Form W was used for the first testing and Form X was used for the second testing.

The first group of the non-advantaged and the first group of the advantaged students were administered the test by tape on the first Wednesday. The second group of non-advantaged and the second group of advantaged students were given the test by regular classroom procedure on the same day. Because of absenteeism, the groups were reduced. The first and second groups of the non-advantaged readers numbered 26 and 27 respectively. The first and second groups of the advantaged numbered 29 and 27.

The following week the test administration procedure was reversed in the groups. There were four testing sessions

each week, these being supervised by the writer. At no time did the classroom teacher take part in the testing procedure.

The T-score formula was used to determine difference of test administration of the non-advantaged groups and the advantaged groups.

The second form of the test, Form X, was administered a week later than the first form, Form W. In order to balance out a carry-over effect from prior experience, the T-score formula was used to determine if there was any significant difference in the taped procedure of administration and the regular procedure of administration when using the combined scores of the taped groups as opposed to the combined scores of the groups using the regular administration.

Analysis of the Data

T-scores of the data of the first and second groups of non-advantaged reading students are presented in Table I. In the first group, significance of the T-scores at the .05 level fell at 2.060, at the .01 level, 2.787. This group yielded the T-score of .52073, indicating no significant difference in the scores of the two test administrations. For the second group, the significance of the T-score at the .05 level fell at 2.056, at the .01 level, 2.779. The second group yielded a T-score of .21896 indicating no significant difference in the scores of the two test administrations.

TABLE I
T-SCORES OF NON-ADVANTAGED READERS

	1st Week	2nd Week	N	t	.05	.01
Group 1	M _r 4.42 S.D. .914	M _t 4.55 S.D. .842	26	.520	2.06*	2.79*
Group 2	M _r 4.61 S.D. 1.12	M _t 4.54 S.D. 1.35	27	.218	2.05*	2.78*

*T-scores for .05 and .01 levels of significance.

On the basis of the first hypothesis that there is no significant difference in the scores of non-advantaged reading students when administered a reading test by Manual directions as opposed to administration by tape, H_{01} is accepted.

The T-scores of the groups of advantaged readers are presented in Table II. In the first group significant difference of the T-scores at the .05 level fell at 2.045, at the .01 level at 2.756. The T-score for this group fell at .60762 indicating no significant difference of the scores of the two test administrations.

Significant difference of the T-scores of the second group at the .05 level fell at 2.052, at the .01 level at 2.771. The second group's T-score fell at the level of .25674, indicating no significant difference of the score on the two test administrations.

On the basis of the second hypothesis that there is no significant difference of scores of advantaged reading students when administered a reading test by Manual directions as opposed to administration by tape, H_{02} is accepted.

The combined scores of all non-advantaged reading students, and the combined scores of all advantaged reading students are presented in Table III. Also in this table are the combined scores of the non-advantaged and advantaged readers.

The T-score of the combined scores of the non-advantaged readers at the .05 level fell at 2.008, at the .01 level,

TABLE II
T-SCORES OF ADVANTAGED READERS

	1st Week	2nd Week	N	t	.05	.01
Group 1	M _r 6.77 S.D. 1.40	M _t 6.54 S.D. 1.52	29	.607	2.04*	2.75*
Group 2	M _r 6.51 S.D. 1.35	M _t 6.41 S.D. 1.29	27	.256	2.05*	2.77*

*T-scores for .05 and .01 levels of significance.

TABLE III
T-SCORES OF COMBINED GROUPS

	1st Week	2nd Week	N	t	.05	.01
Non-ad.	M_r 4.52 S.D. 1.02	M_t 4.54 S.D. 1.12	53	.117	2.00*	2.67*
Advan.	M_r 6.35 S.D. 1.45	M_t 6.49 S.D. 1.32	56	.535	2.00*	2.66*
Total Groups	M_r 5.60 S.D. 1.64	M_t 5.54 S.D. 1.56	109	.278	1.98*	2.62*

*T-scores of .05 and .01 levels of significance.

2.678. The T-score for the non-advantaged reader's score fell at the .11799 level, indicating no significant difference at either level.

The T-scores of the combined scores of the advantaged readers at the .05 level fell at 2.004, at the .01 level at 2.669. This group's T-score fell at .53561, indicating no significant difference at either level.

The T-scores of the combined scores of the non-advantaged and advantaged reading students are presented in the last section of Table III. At the .05 level the significant difference of the T-score fell at 1.984, at the .01 level at 2.626. T-score for these combined groups fell at the .278 level indicating no significant difference at the .05 level or the .01 level.

Although there was no hypotheses given for the combined scores of the groups of the non-advantaged readers and the advantaged readers, the means of the combined groups were calculated by the T-score formula. There was no significant difference of the means of the combined groups.

This chapter presented the results of the data of selected groups of non-advantaged and advantaged reading students when being assessed by different test administration procedures.

The mean scores of the four groups were calculated using the T-score formula to test two hypotheses:

H_{01} stated that there would be no significant differences in the scores of non-advantaged reading students

when administered a reading test by Manual directions as opposed to administration by tape. H_{01} was accepted.

H_{02} stated that there would be no significant difference in the scores of advantaged reading students when administered a reading test by Manual directions as opposed to administration by tape. H_{02} was accepted.

The combined scores of the non-advantaged and advantaged readers, and the combined scores of the total groups were calculated using the T-score formula. There was no significant difference of the means of the combined groups.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

This chapter will be presented in three sections. First, a general summary of the investigation will be given. The second section will be concerned with the conclusions drawn from the study. The last section will discuss recommendations for further research.

Summary

The purpose of this study was to determine if different procedures of administration of standardized reading tests may be desirable when such tests are to be given to non-advantaged reading students.

The data utilized in this investigation for the hypotheses were collected from 53 non-advantaged reading students divided into two groups, and from 56 advantaged reading students divided into two groups. They were administered the Stanford Reading Test, Form W, in April, 1973, and the Stanford Reading Test, Form X, a week later, in April, 1973.

The non-advantaged reading students were identified by using Bond and Tinker's (1968) formula, years in school multiplied by I.Q., plus one. This formula determined the students' expectancy level.

The Stanford Reading scores from the Stanford Achievement Tests and the Otis I.Q. scores of the fifth and sixth grade students were used. These were on record in the school files from the year 1972. The students expectancy score was subtracted from his score on the reading section of the Stanford Achievement Tests. Those students who had a score falling one and one-half years or more below their reading expectancy score were considered non-advantaged readers for the study. Those falling one and one-half years or more above their reading expectancy were considered advantaged readers for the study.

The names of the identified non-advantaged students were alphabetically arranged and divided into two equal groups of 30 by random sampling. Proof of random sampling was obtained by using the T-score formula to determine if there were significant differences in the mean of the reading scores. The same procedure was used to select two groups of 30 from the identified advantaged readers group. Because of absenteeism the non-advantaged groups were reduced to 26 and 27. The advantaged groups were reduced to 27 and 29 for the same reason.

The four groups were given the Stanford Reading Tests on two successive Wednesdays in April, 1973. Form W was used for the first testing and Form X for the second testing.

One group of the non-advantaged and one group of the advantaged readers were administered the test by tape on the first Wednesday and the other groups of non-advantaged and

advantaged students were given the test by regular classroom procedure. The following week the test procedure was reversed in the groups. There were four testing sessions each week, supervised by the writer.

The second form of the test, Form X, was administered a week after the first form, Form W. In order to balance out a possible carry over effect from prior experience, the T-score formula was used to determine if there were any significant differences in the score of the taped procedure and the regular procedure of administration when using the combined scores of the taped groups as opposed to the combined scores of the groups using the regular administration.

The T-score formula was used to determine the significant difference of test administration of both the non-advantaged and the advantaged students.

The testing was conducted in a classroom especially arranged by the school personnel. The test was not pre-announced to the students.

The instruments used in this study were the Stanford Reading Tests, Intermediate II, Forms W and X.

The tapes were prepared by the writer. The instructions were read onto two recorders and the tapes were edited for errors. There were no errors and the recorder did not fail during administration.

The T-score formula was used to determine the significant difference of the scores at the .05 level and the .01 level of both the non-advantaged and the advantaged students.

Two hypotheses were stated in the null concerning the scores of the non-advantaged reading students and the scores of the advantaged reading students.

The hypotheses are as follows:

Ho₁: There is no significant difference in the scores of the non-advantaged reading students when administered a reading test by manual directions as opposed to administration by tape.

Ho₂: There is no significant difference in the scores of advantaged reading students when administered a reading test by manual directions as opposed to administration by tape.

Conclusions

The results of the analysis of data in the present investigation warrant the following conclusions:

1. Scores of non-advantaged reading students were not significantly different when administered the directions of the test by manual directions as opposed to tape administration.
2. Scores of advantaged reading students may be significantly different at the .05 level when administered the directions of the test by manual directions as opposed to tape administration.

Limitations

1. The reliability of the measuring instruments will to some degree affect the reliability of any conclusion drawn on this study.
2. Other than the students being identified as non-advantaged and advantaged reading students of the fifth and sixth grades of the Pawhuska Elementary School, Pawhuska, Oklahoma, there was no attempt to control for other student differences.

Recommendations

The present study has made a contribution to reading test administration to non-advantaged reading students by elimination of one method of administration--taped directions. Additional research is needed to seek better methods of testing non-advantaged reading students. Recommendations for further research based on the present study are as follows:

1. The results of this study were based on data collected from the fifth and sixth grades of a small school. It is recommended that a similar study be conducted with larger and different, urban and rural, populations.
2. Further research is needed to identify better procedures of administering test material to non-advantaged reading students at these grade levels.

3. Research is needed to determine if the taped procedure of test administration might be more effective at lower and higher grade levels.

Concluding Statement

The present investigation was designed to study two procedures of reading test administration to non-advantaged reading students. It was found that there is no significant difference in the scores of the two procedures.

BIBLIOGRAPHY

- Bereiter, V., and S. Engelmann. Teaching Disadvantaged Children in the Preschool. Englewood Cliffs, N. J.: Prentice-Hall, 1966.
- Bereiter, C., and S. Engelmann. "Teaching Disadvantaged Children the Language of Instruction." Canadian Education and Research Digest, Vol. 8 (October, 1967), pp. 126-36.
- Bond, G. L., and J. A. Tinker. Reading Difficulties: Their Diagnosis and Correction. New York: Appleton-Century-Crofts, 1967.
- Bromwich, Rose. "The Young Child's Language and Effective Learning." Paper delivered at N.C.T.E. National convention, Las Vegas, 1971.
- Briggs, Thomas H., and George H. Armacost. "Results of an Oral True-False Test." Journal of Educational Research, XXVI (April, 1933), pp. 595-96.
- Carmichael, Stokely, and Charles V. Hamilton. Black Power. Chicago, Ill.: Random House, Inc., 1967.
- Cronbach, Lee J. Essentials of Psychological Testing. New York: Harper and Row, 1960.
- Farb, Peter. "The American Indian: A Portrait in Limbo." Saturday Review, October, 1967, pp. 165-172.
- Farnham-Diggory, S. Cognitive Synthesis in Negro and White Children. Chico, California: Monograph 35. Chico State College, Monograph Society for Research in Child Development, 1970, pp. 78-80.
- Friedman, Norman L. "Cultural Deprivation: A Commentary in the Sociology of Knowledge." Journal of Educational Thought, Vol. I (August, 1967), pp. 88-89.
- Frost, Joe L., and Glenn R. Hawkes, The Disadvantaged Child. New York: Houghton Mifflin Company, 1970.
- Frostig, Marianne. "Visual Perception, Integrative Functions and Academic Learning." Journal of Learning Disabilities, Vol. 5, No. I (January, 1972), pp. 5-17.

- Horn, Thomas D. "Listening-Speaking Problems of Disadvantaged Children." Reading: The Third Level. Division of Summer Sessions, Syracuse, New York: Syracuse University, 1968, pp. 13-20.
- Johnson, F. Craig, and Kenneth Frandsen, "Administering the Brown-Carlsen Listening Comprehension Test." Journal of Communication, Vol. 13 (March, 1963), pp. 38-45.
- Jones, John Alfred. "An Investigation of Scores on the Brown-Carlsen Listening Test as Correlated With Subjective Instructor Ratings and as Influenced by Media of Presentation." (Master's Thesis, University of Washington, Seattle, 1961).
- Knox, Lawrence B., and Verner M. Sims. "The Reliability and Validity of Multiple Response Tests When Presented Orally." Journal of Educational Psychology, XXIII (December, 1932), pp. 656-662.
- Noll, V. H. Introduction to Educational Measurement. Boston: Houghton Mifflin, 1957.
- Odom, Chalres L., and Ray W. Miles. "Oral Versus Visual Presentation of True-False Achievement Tests in the First Course of Psychology." Educational and Psychological Measurement, XI (Autumn, 1951), pp. 470-477.
- Schrag, Peter. "Appalachia: Again the Forgotten Land." Saturday Review, January 27, 1968, pp. 14-18.
- Stumpf, Franklin N. "Listening Versus Reading Method in True-False Examinations." Journal of Applied Psychology, XV (1931), pp. 555-562.
- Stumpf, Franklin N. "Oral Versus Printed Method in the Presentation of the True-False Examination." Journal of Educational Research, XVIII (December, 1928), pp. 423-424.
- Tobin, Richard L. "One Million Migrants." Saturday Review, August 17, 1968, pp. 12-15, 61-63.
- Utigard, Carl N. "A Comparison of the Effectiveness of Magnetic Tape and the Regular Classroom Teacher Upon Growth in Listening Ability." (Master's Thesis, University of Washington, Seattle, 1962).
- Westover, Frederick L. "A Comparison of Listening and Reading as a Means of Testing." Journal of Educational Research, Vol. 52, No. 1 (September, 1958), pp. 23-27.

Wyatt, Woodrow, W., and Charles M. Bridges. Statistics for the Behavioral Sciences. Boston: D.C. Heath and Company, 1967.

APPENDIX A

MIDDLE OF GRADE 5

Test	r ² _{II}	r ³ _{KR 20}	St. Error ⁴ Measurement
Word Meaning	.89	.89	4.5
Para. Meaning	.93	.92	4.0

MIDDLE OF GRADE 6

Test	r ² _{II}	r ³ _{KR 20}	St. Error ⁴ Measurement
Word Meaning	.90	.90	4.5
Para. Meaning	.93	.92	5.0

1. Values reported are based on a sample of 1,000 cases from each grade, 5.6 and 6.6 drawn randomly from 76 school systems testing in all grades 1-8 in national standardization.
2. Split-half reliability coefficients corrected by the Spearman-Brown Prophecy Formula.
3. Estimate of Kuder-Richardson Formula 20 using Saupe's Formula.
4. Standard error of measurement in terms of grade scores.

APPENDIX B

T-SCORE FORMULA

$$T = \frac{M_1 - M_2}{\frac{\Sigma X_1^2 + \Sigma X_2^2}{N_1 + N_2 - 2}} \cdot \frac{N_1 + N_2}{N_1 N_2}$$

VITA

Margaret Shaffer Atkin

Candidate for the Degree of
Doctor of Education

Thesis: AN INVESTIGATION OF INSTRUCTION TO READING ACHIEVEMENT MEASUREMENT FOR NON-ADVANTAGED READING STUDENTS IN GRADES 5-6

Major Field: Student Personnel and Guidance

Biographical:

Personal Data: Born in Independence, Kansas, February 22, 1918, the daughter of Fred J. and Myrtle Kahle Shaffer.

Education: Attended public schools in Arkansas City, Kansas. Graduated from Arkansas City High School in 1933. Attended Arkansas City Junior College in 1934 and 1962; received an Associate of Arts degree from Arkansas City Junior College in 1962; received a Bachelor of Arts degree in elementary education from Southwestern College, Winfield, Kansas, in 1963; received a Master of Science degree in Student Personnel and Guidance from Oklahoma State University, Stillwater, Oklahoma, in 1967; completed requirements for the Degree of Doctor of Education in July, 1975.

Professional Experience: Fifth grade teacher, 1963-1965, at Dexter Public Schools, Dexter, Kansas; counselor-teacher 1965-1966, at Cedar Vale Public Schools, Cedar Vale, Kansas; counselor on a federal program headquartered at Pawhuska, Oklahoma, 1966-1967; counselor-coordinator, 1967-1968, on above program; counselor-teacher, 1968-1969, at Shidler Public Schools, Shidler, Oklahoma; counselor, 1969-1970, at Arkansas City Junior High, Arkansas City, Kansas; trainee, January to June, 1971, Child Psychiatry Department of Oklahoma University Medical School, Oklahoma City, Oklahoma; school psychologist, 1971-1975, Pawhuska Public Schools, Pawhuska, Oklahoma.