

TEACHER EXPECTATIONS: DETERMINANTS  
OF PUPILS' READING ACHIEVEMENT

By

DAVID V. ELIJAH, JR.

Bachelor of Arts  
Purdue University  
West Lafayette, Indiana  
1966

Master of Science  
Oklahoma State University  
Stillwater, Oklahoma  
1972

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  
May, 1974

Thesis  
1974 D  
E 42t  
Cop. 2

MAR 13 1975

TEACHER EXPECTATIONS: DETERMINANTS  
OF PUPILS' READING ACHIEVEMENT

Thesis Approved:

Darel D. Ray  
Thesis Adviser

Leslie Olson

Bill F. Elsom

Paul Gardner

Larry M. Perkins

D. N. Denton  
Dean of the Graduate College

902068

## ACKNOWLEDGEMENTS

The writer wishes to express his deep appreciation to the chairman of his Advisory Committee, Dr. Darrel D. Ray, for his constant help and understanding attention to the problems encountered in this study. Thanks are also extended to Dr. Billy F. Elsom, Dr. Russell L. Dobson, Dr. Paul G. Warden, and Dr. Larry Perkins for their assistance in the formulation and preparation of this study.

Sincere appreciation is extended to Dr. Anne Bell and Mr. Frank Borovetz for helping to facilitate the testing for this study. Grateful acknowledgement is given to the principals and teachers of the Muskogee, Oklahoma schools for their cooperation in this study.

My sincerest appreciation is expressed to my colleagues and friends at Oklahoma State University for their help during the collection of the data for this study. Without their help and cooperation this study could not have been conducted.

This study is dedicated to my parents, Mr. and Mrs. Victor Elijah.

## TABLE OF CONTENTS

Chapter	Page
I. PRESENTATION OF THE PROBLEM . . . . .	1
Introduction . . . . .	1
Need for the Study . . . . .	2
Statement of the Problem . . . . .	4
Purpose of the Study . . . . .	4
Hypotheses . . . . .	5
Definitions of Terms . . . . .	5
Scope of the Study . . . . .	7
Limitations of the Study . . . . .	8
Assumptions of the Study . . . . .	8
Methodology. . . . .	9
II. REVIEW OF THE LITERATURE. . . . .	10
Introduction . . . . .	10
Teacher Expectations as Determinants of Pupil Achievement . . . . .	11
Socioeconomic Status as a Determinant of Pupil Achievement . . . . .	21
Teacher Judgement of Reading Readiness . .	30
Summary. . . . .	33
III. DESIGN AND METHODOLOGY. . . . .	35
Introduction . . . . .	35
Description of the Population. . . . .	35
Testing Procedures . . . . .	36
Instruments Used in the Study. . . . .	37
Statistical Design . . . . .	42
Summary. . . . .	43
IV. TREATMENT OF DATA AND ANALYSIS OF RESULTS . . .	44
Tests of the Hypotheses. . . . .	45
Summary. . . . .	53

Chapter	Page
V. SUMMARY AND CONCLUSIONS . . . . .	55
General Summary of the Investigation . . .	55
Conclusions. . . . .	57
Recommendations. . . . .	59
BIBLIOGRAPHY . . . . .	60

LIST OF TABLES

Table	Page
I. Analysis of Variance of Difference Between Word Reading for Different Socioeconomic Status and Falsified Metropolitan Readiness Test Scores Two Months After Data Presented to Teachers . . . . .	46
II. Analysis of Variance of Difference Between Paragraph Meaning for Different Socioeconomic Status and Falsified Metropolitan Readiness Test Scores Two Months After Data Presented to Teachers . . . . .	46
III. Analysis of Variance of Difference Between Vocabulary for Different Socioeconomic Status and Falsified Metropolitan Readiness Tests Scores Two Months After Data Presented to Teachers . . . . .	47
IV. Analysis of Variance of Difference Between Word Reading for Different Socioeconomic Status and Falsified Metropolitan Readiness Tests Scores Six Months After Data Presented to Teachers . . . . .	47
V. Analysis of Variance of Difference Between Paragraph Meaning for Different Socioeconomic Status and Falsified Metropolitan Readiness Tests Scores Six Months After Data Presented to Teachers . . . . .	48
VI. Analysis of Variance of Difference Between Vocabulary for Different Socioeconomic Status and Falsified Metropolitan Readiness Tests Scores Six Months After Data Presented to Teachers . . . . .	48
VII. Spearman Rank Order Correlations of Metropolitan Readiness Test Rankings and Pretest Teacher Rankings of Reading Readiness . . . . .	50

Table	Page
VIII. Spearman Rank Order Correlations of Pretest Teacher Rankings of Reading Readiness and Posttest Teacher Rankings of Reading Readiness . . . . .	52
IX. Spearman Rank Order Correlations of Metropolitan Readiness Test Rankings and Posttest Teacher Rankings of Reading Readiness . . . . .	52



## CHAPTER I

### PRESENTATION OF THE PROBLEM

#### Introduction

The purpose of a school is to provide opportunities for all children to learn skills necessary to function in our society. Success in reading is a skill highly valued in our culture. Without the ability to read, a person is very limited in the choices he can make throughout his life.

School is important to a learner. It is the institution responsible for developing the many cognitive skills necessary for achievement in our society. It is the setting in which the learner spends a great amount of time. During the time he is in school, most of his attitudes toward learning are developed and reinforced. During the early years in school, the learner can develop a sense of industry and good task identification if he is successful. It is, therefore, very important that this time in a child's life be productive and he be aware of his cognitive successes in school (Erikson, 1968).

A teacher's attitudes and beliefs about children will make a difference in her classroom performance. Teachers, who believe that children can learn, will try to teach them.

Conversely, teachers who feel that certain children cannot learn will go through the motions of teaching, while not expecting much achievement from the children. The teacher's beliefs about her students can be a limiting factor in determining how well the students perform in school (ASCD, 1962).

The primary grades of school are very important grades in establishing the child's academic self-image (Erikson, 1968). The self-image developed may be one of success or one of failure. A child's success in school is mainly determined by his success in learning to read. If a student has difficulty in learning to read, he may look upon himself as a failure (Bond & Tinker, 1967).

#### Need for the Study

Many people now believe that a teacher's expectations for a child's success can be a determining factor as to whether or not the child does indeed successfully learn in school (Rosenthal, 1971). If a teacher's expectations for a child's success can be controlled, then the child's achievement can be better assured.

The first grade teacher has very little concrete information on which to base her academic expectations for her pupils. One of the most important tools available to the first grade teacher is a reading readiness test. This gives the teacher a major indication of how easily she may expect the child to learn to read. The teacher's interpretation of reading readiness test results may largely influence her

expectations for her pupils' success. These expectations in turn help determine the teacher's attitude towards the child. Therefore, the influence of reading readiness measures may be extensive.

Another indicator on which a teacher bases her academic expectations for her children is the socioeconomic status of the children in her classroom. She is aware of the children's backgrounds and has expectations for her children's achievement based upon her beliefs about socioeconomic status (Erikson, 1968). These beliefs about socioeconomic status and their relationship to achievement can alter her expectations for her children's level of reading achievement. The teacher has differing expectations dependent on the socioeconomic status of the child.

The classroom teacher has an opinion of the reading readiness status of the pupils in her classroom. To what degree the teacher's opinion of reading readiness status is influenced by the socioeconomic status of her pupils needs to be examined. The teacher's opinions of her pupils' reading readiness status also need to be examined to determine if the teacher is a reliable judge of reading readiness status. A better understanding is needed of the expectations that teachers hold for children's failures in learning to read. It is also necessary to see if teacher expectations can be controlled.

The purpose of the study was to ascertain to what extent teacher expectations influence reading performance of

students at the first-grade level. This study was designed to investigate whether student's reading achievement can be increased by influencing the teacher's expectations for a student through the falsification of the student's reading readiness scores.

Answers to the following questions were sought:

1. Does providing the teacher with falsified reading readiness scores influence reading achievement of students?
2. Can teacher expectations be influenced by providing the teachers with falsified reading readiness scores?
3. Are teacher rankings of readiness status as reliable as the Metropolitan Readiness Tests rating?

#### X Statement of the Problem

The problem concerning this study is that it is unknown if teacher expectations do influence a child's reading performance, or, if teacher expectations do influence a child's reading performance, to what extent teacher expectations may influence reading performance.

#### Purpose of the Study

The purpose of this study was to determine to what extent teacher expectations influence reading performance of students. Another purpose of this study was to determine if teacher expectations can be manipulated.

## Hypotheses

The hypotheses to be tested are stated in the null form as:

1. There is no significant difference in reading achievement between groups with different levels of reading readiness who are perceived by the teacher as being at the same state of reading readiness performance.

2. Student socioeconomic status and teacher knowledge of student reading readiness performance will not interact to shape reading achievement.

3. There is no significant relationship between a teacher ranking of reading readiness status of her children before readiness tests have been administered and a ranking of the children's Metropolitan Readiness Test scores.

4. There is no significant relationship between a teacher ranking of reading readiness status of her students before readiness tests have been administered and teacher ranking of reading readiness status after reading readiness tests have been administered.

5. There is no significant relationship between a teacher ranking of reading readiness status of her students after readiness tests have been administered and a ranking of the children's Metropolitan Readiness Test scores.

## Definitions of Terms

For the purpose of this study the following definitions will be used.

1. Reading Readiness will be defined in terms of total scores on the Metropolitan Readiness Tests, Form B, as established by the norms for that instrument.

2. Socioeconomic Status of the pupil will be determined by the head of the household's (parent's) occupation as ranked by the Otis Dudley Duncan Socioeconomic Index for Occupations based on the North-Hatt Scale. A high socioeconomic status child in this study is one whose parent's occupation ranks in the top third of his class by the Otis Dudley Duncan Socioeconomic Index for Occupations. A low socioeconomic status child will be defined as a child whose parent's occupation ranks in the bottom third of his class on the Otis Dudley Duncan Socioeconomic Index for Occupations.

3. The North-Hatt Scale is a social ranking scale of occupations covering the range from U. S. Supreme Court Justice to shoe shiner. The scale was developed under the direction of Cecil North and Paul Hatt in the National Opinion Research Center at Ohio State University.

4. Normal Range of Intelligence is defined as a performance score of between 84 and 116, inclusive, on the Short Test of Educational Abilities, Level I. These scores are within one standard deviation of the mean of the instrument's normative population. The Short Test of Educational Abilities, Level I, is a standardized group intelligence test for kindergarten and first-grade children.

5. First-Grade Children are those children who are enrolled in grade one for the first time and are not repeating the grade at the time of initial testing.

6. Teacher Expectations are defined as a change in the teacher's reading readiness rankings as a result of having been presented with reading readiness test results.

#### Scope of the Study

This study includes an analysis of the reading achievement scores made by low reading readiness first-grade pupils, some of whom had false reading readiness scores reported to their teachers. Comparisons of achievement scores of the control groups with the children for whom false scores were reported were made twice during the first grade. Comparisons were made between high and low socioeconomic status children. A comparison was made between teacher rankings of reading readiness before and after reading readiness test scores were reported. A comparison was made between teacher rankings of reading readiness status and the Metropolitan Readiness Test ranking.

The subjects in this study included all of the first-grade children in one northeastern Oklahoma city who were average in intelligence, of either high or low socioeconomic status, and who received a readiness score of minus one standard deviation calculated by individual classroom on the Metropolitan Readiness Tests. The final sample included

sixty-four children from 26 classrooms in 12 elementary schools in the city.

#### Limitations of the Study

This study is limited to first-grade pupils from one city in northeastern Oklahoma.

The tests used in this study were only a sample of the measures that might have been used. Other tests might have yielded different results.

#### Assumptions of the Study

1. The random sampling procedure supports the assumption that the children selected were representative of the population of all first-grade classrooms. There are no known reasons for believing that this school year and these first-grade children were not typical of a much larger population of first-grade classrooms. Strictly considered, however, the conclusions of this study can only be generalized to the particular population sampled.

2. The Otis Dudley Duncan Socioeconomic Index for Occupations is a reliable instrument for determining the socioeconomic status of pupils.

3. The Short Test of Educational Abilities, Level I, is a reliable measure of intelligence.

4. The Metropolitan Readiness Test, Form B, is a reliable measure of reading readiness skills.



5. The Stanford Achievement Tests, Primary I Battery, Forms W and X, are an adequate measure of reading achievement.

#### Methodology

A factorial analysis of variance was used to test for significant differences in achievement between the experimental groups. The teacher rankings of reading readiness were analyzed using the Spearman Rank-Order Correlation.

## CHAPTER II

### REVIEW OF THE LITERATURE

#### Introduction

The area representing information relevant to the problem discussed and examined in this paper is primarily confined to a review of the literature concerning the role of a teacher's expectations as a determinant of pupil achievement. The studies reviewed are those that demonstrate the phenomenon of teacher expectations as determinants of pupil achievement and those that fail to demonstrate the phenomenon of teacher expectations as determinants of pupil achievement. Studies relative to the relationship existing between socioeconomic status and reading achievement are reviewed in a second section. A third section in the review of the literature concerns teachers' ability to appraise their students' level of reading readiness through informal measures. The predictive value of the Metropolitan Readiness Tests as an indicator of reading achievement in the primary grades has been well established in studies by Karlan (1957), Simpson (1960), and Akers (1967).

Teacher Expectations as Determinants  
of Pupil Achievement

One of the earliest studies conducted to test the hypothesis that teacher expectations affect pupil performance was done by Rosenthal and Jacobson (1966). Rosenthal and Jacobson hypothesized that teacher expectations would significantly influence pupil's I. Q. gains. All of the children in a low socioeconomic elementary school were used in their study. There were 18 classrooms in the elementary school, three at each of the six grade levels. Heterogeneous ability grouping was used throughout the school. All children in the school were administered a nonverbal intelligence test, the Tests of General Ability or TOGA. Approximately 20 percent of the children from each classroom were randomly chosen to be the experimental group. Each teacher was given the names of these children in her room and was told that test results indicated that those children could be expected to show remarkable gains in intellectual competence during the remainder of the school year. Eight months later all children in the school were again retested with TOGA and a gain score was computed for each child. A t test was used to test for significance. The results are as follows:

1. As a whole, those children from whom the teachers had been led to expect greater intellectual gain showed significantly greater gains in I. Q. scores than did the control group children ( $p = .02$ , one-tailed test).
2. The effects of teachers' expectancies were not uniform across the six

different grade levels. The lower the grade level, the greater was the effect with the effects being the greatest in the first and second grades. Significance for the first grade was  $p = .002$  and second grade significance was  $p = .02$ . Differences between grades 3, 4, 5, and 6 were not significant. Rosenthal and Jacobson give a number of conclusions as to why teachers' expectancy effects operated primarily at the lower grade levels. They are:

1. Younger children have less well-established reputations so that the creation of expectations about their performance would be more credible.
2. Younger children may be more susceptible to the unintended social influence exerted by the expectations of their teacher.
3. Younger children may be more recent arrivals in the school's neighborhood and may differ from the older children in characteristics other than age.
4. Teachers of lower grades may differ from teachers of higher grades on a variety of dimensions which are correlated with the effectiveness of the unintentional communication of expectancies (Rosenthal & Jacobson, 1966).

Jose's (1970) study was a partial replication of Rosenthal and Jacobson (1966). The basic concern of his investigation was to determine if students, who had been identified to teachers as being capable of doing better work, would show a greater increase in intellectual growth and school achievement than would a control group. His sample consisted of 18 teachers, nine first grade and nine second grade, and 144 students, eight from each classroom. The teachers were randomly selected and randomly assigned to

experimental and control groups. The students were pre-tested with TOGA and the subjects in the experimental group were presented to their teachers as "late bloomers." The experimental treatment resulted in little difference in the performance of the experimental and control subjects. However, Jose concluded that there were many similarities in results between his study and Rosenthal's (1966). Some students achieved high gain scores in I. Q. in both studies.

Goldsmith and Fry (1971) tested the question: Will a high expectancy prediction make a significant difference in I. Q. or reading achievement of 10th grade students. They randomly selected 112 tenth grade students as a control group from a middle-class high school in New Jersey. Pre-testing and posttesting was done using the Sequential Tests of Educational Progress and TOGA. Teachers were given a bogus high expectancy prediction for the 112 experimental children. Their results showed no significant difference between the two groups in either I. Q. or reading achievement gains. Goldsmith and Fry concluded that within the limits of their study, the efficacy of the teacher expectancy effect on the high school level failed to be supported.

Beez (1968) was also concerned with changes in teaching behavior and performance by pupils. He also attempted to investigate the question of how teacher's expectancies become translated into behavior so as to elicit the expected pupil behavior. Beez worked with 60 children from the summer Headstart Program in Bloomington, Indiana. The children

ranged in age from 5 years, 7 months to 6 years, 6 months, and had I. Q.'s on the Peabody Picture Vocabulary Test from fifty-five to 127, with a mean of 91. Children were randomly assigned to either a "low ability" or a "high ability" group. Sixty graduate students in the School of Education of Indiana University during the summer of 1967 served as the subjects of the study. The subjects were given a folder containing a faked "evaluation" of the child. All reports for each group were identical except for name and age. The "low ability" children were described negatively and the "high ability" children were described in positive terms. Each subject taught one child a specific task. There was a significant difference ( $p = .001$ ) in the achievement of "low ability" children and "high ability" children. Beez concluded that his results strongly supported the findings of Rosenthal and Jacobson (1966) that pupils are influenced by their teachers' expectancies and have a tendency to behave accordingly.

Conn, Edwards, Rosenthal, and Crowne (1968), doing research designed to study the accuracy with which the child perceived vocal expression of emotion as a factor of teacher-child interaction, closely replicated Rosenthal and Jacobson (1966). They administered the Tests of General Ability to students in grades one through six of an upper middle-class elementary school. The instrument was disguised as a test designed to predict academic "blossoming" or intellectual gain. Approximately 23 percent of the students

were randomly assigned to the experimental group. The names of these children in the experimental group were given to their teachers and told that they would show unusual intellectual gains during the coming year. The students were re-tested four months later and again a full year after the children had left the teachers who had been given special expectations for the children in the experimental group. The researchers (Conn et al., 1968) obtained results parallel to those of Rosenthal and Jacobson (1966). The experimental group showed a tendency to gain more in I. Q. than the control group for one semester. There were no significant differences in the groups over three semesters. They concluded that positive expectations do not necessarily lead to positive results for all pupils.

In an experiment conducted at the USAF Academy Preparatory School, Schrank (1968), tested the null hypothesis that randomly grouped classes will show no differences in academic achievement when ability grouping is simulated and classes are labeled with ability level designations. Schrank used a table of random numbers to assign 100 airmen to one of five class sections. Mathematics was the subject matter used and the experiment lasted for seven months. Ability-level labels in the form of numerical designations were assigned to the sections. Neither the students or the instructors knew the grouping was randomized. The academic grades of the students served as the independent variable. Schrank found that the difference between the means of the

highest-labeled section and the lowest-labeled section to be significant at the 0.1+ percent level. It was found that the higher-level labeled section achieved a higher mean than the next lower labeled section. Schrank concluded that there is definitely a labeling effect present in simulated ability grouping even though the grouping is actually random.

In a second experiment, Schrank (1970), replicated his first experiment except that he informed instructors that the ability grouping was only being simulated. The results of this experiment failed to demonstrate any significant difference between the different groups. He concluded that the labeling effect was not a predominant factor in this study. He also concluded that the teacher plays a major role in the production of the labeling effect and that it is the teacher's reaction to his own expectation of the pupil's performance that is reflected in the teacher's grading standards and teaching methods thus creating the performance the teacher expects.

Clairborn (1969) tried to replicate Rosenthal and Jacobson (1966) in his work with first-grade children. His first purpose was to quantify some in-class teacher-pupil behavior in an attempt to isolate the variables that might cause the I. Q. gains in children for whom higher expectations were held. Clairborn's second purpose was to replicate Rosenthal and Jacobson (1966). His sample consisted of three first-grade classrooms from each of three middle-



class suburbs in upstate New York. Classes were assigned to the "Bias" and "No Bias" condition at random and approximately 20 percent of the children were designated as "potential intellectual bloomers" to their teachers. A pretest and posttest difference in I. Q. as measured by the TOGA was used to assess I. Q. gains. The analysis of covariance yielded an F ratio of 2.12 (df - 1/101) which was not significant. He also found no significant differences when I. Q. subtest scores were compared. There was weak evidence of a differential change in the teacher-pupil interactions with the non-special pupils in the "Bias" classrooms when compared with the non-special pupil in the "No Bias" classrooms ( $F = 3.6$ ,  $df = 8/95$ ,  $p = .01$ ). He concluded that further research needs to be conducted before the conclusions of the Rosenthal and Jacobson (1966) experiments become accepted as psychological fact.

Palardy (1968) also used the rationale of the self-fulfilling prophecy to investigate the effect of teacher's beliefs on pupils' achievement. His main purpose was to determine whether teachers' reported beliefs about first-grade boys' probable success in reading had any significant effect on the measured achievement in reading that the pupils in their classes attained. Palardy was particularly interested in the effect of these beliefs on the boys' achievement. Sixty-three first-grade teachers in an Ohio city completed a questionnaire designed to elicit from the teachers a report of their beliefs regarding the probable success

of first-grade boys in learning to read. Of the teachers who responded to the questionnaire, ten teachers were selected to be in the study. Five of these teachers (Group A) believed the boys would be as successful as girls in learning to read. Five of the teachers (Group B) did not believe boys would be as successful as girls in learning to read. In May all students in the ten classrooms were given reading achievement tests. The combined effect of pupils' sex and teachers' beliefs resulted in a lower achievement score for the boys in Group B, those boys whose teachers believed that first-grade boys are far less successful than girls in learning to read. The two groups of boys' scores were compared by an F test. The comparison revealed a difference in mean scores that closely approximated significance, favoring the boys in Group A ( $F = 3.174, p = .08$ ).

Palardy (1968) concluded that when first grade teachers reported that they believed that boys are far less successful than girls in learning to read, the boys of those teachers did achieve less well than boy pupils whose teachers reported they believed that boys are as successful as girls in learning to read. He also concluded that in terms of the self-fulfilling prophecy it can be stated that when teachers believed that boys are not as successful as girls in learning to read, the boys in their classes are far less successful than the girls and, conversely, when teachers believed that boys are as successful as girls, the boys are as successful as girls.

Brophy and Good (1970) extended previous studies of the effects of teacher expectations on children's academic performance by studying how children change so as to begin to conform to teacher expectations. They approached the problem through classroom interaction analysis. Their subjects were four first-grade classrooms in a small rural Texas school district. The teachers ranked the children in order of their achievement. These rankings were used as an indicator of the teachers' expectations for classroom performance. In each classroom, three boys and three girls low on the teacher's list were chosen for observation. All interaction between these students and the teacher were recorded and categorized by observers. Their results may be summarized in the statement that outside of the reading group at least, the highs seek out the teacher and initiate interactions with her more frequently than the lows, especially in work-related interactions. The teachers called on the highs to answer more questions and more frequently criticized the lows, especially boys. The big difference between the highs and the lows was in quality rather than quantity of interaction with the teacher. Brophy and Good concluded that teachers systematically discriminate in favor of the highs over the lows in demanding and reinforcing quality performance, and the nature of this differential treatment is such as to encourage the children to begin to respond in ways which would confirm teacher expectancies.

In short, they concluded that teacher expectations function as self-fulfilling prophecies.

Good (1970) was also interested in how teacher expectations are transmitted to the pupils. Good studied four first-grade classrooms for the purpose of determining if teachers call on pupils perceived as high achievers significantly more than they call on pupils perceived as low achievers. Teachers who took part in the study were asked to provide the investigator with a seating chart and a list of pupils ranked in order of achievement. Twelve students from each classroom were observed; the first four pupils, the last four pupils, and the middle four pupils on the teachers' achievement list. Good's results showed that opportunities to respond were related to pupil achievement as rated by the teacher. He states that the especially important findings of this study is that low achievers were deprived of opportunities to respond in competitive, non-reading classroom situations. The difference in the number of opportunities to respond as ranked as High, Middle, or Low achievement was significant at the .001 level. He concluded that pupils who have a low achievement record continue to show less achievement than their classmates and that teachers' actions contribute to the vicious circle because teachers "know" that low achievers cannot provide the answer and do not call on them.

The body of research leads one to conclude that the phenomenon of teacher expectations does exist and is an

influential factor in determining how much is learned in the primary school classroom. Although the research demonstrating teacher expectations has been primarily concerned with I. Q. gains, it has direct implications for reading instruction. If I. Q. scores can be increased by giving a teacher false expectations, it is a reasonable supposition to believe that other cognitive pupil gains can be achieved in the same manner.

#### Socioeconomic Status as a Determinant of Pupil Achievement

One of the earliest studies to determine the relationship between ability and social status was conducted by Havighurst and Janke (1944). In this study, 110 ten year old children from a "typical middlewestern community" constituted the sample. Four other ten year olds resided in the community at that time but were not given the entire battery and were excluded from the final sample. Havighurst and Janke utilized the social status method of studying the community and placed the families of the ten year old children on a scale of social status. They found five social classes in the community. Group A, the so-called "top crowd" had no ten year old children. Group B, composed mostly of the families of professional men, had one child in the ten year old group. Group C, the small business men and other white-collar workers, had 23 percent of the ten year olds. Group D, the semi-skilled workers and labors, contained 62

percent of the ten year olds. Group E, the "river rats" and the "bottom of the heap," contributed fourteen percent of the ten year olds.

All of the children were tested separately by one of the authors. The following tests were administered: 1. the Revised Stanford Binet, Form 1, 2. the Cornell-Cox Performance Ability Scale, 3. the Iowa Silent Reading Test, New Edition, Form Am, 4. the Minnesota Paper Form Board, Revised, Form AA, 5. the Minnesota Mechanical Assembly Test (for boys), and the Chicago Assembly Test for Girls, 6. the Porteus Maze Test, and 7. the Goodenough Draw-A-Man Test.

The test results were compared for social class groups by calculating the ratio of the difference of the means to the standard error of that difference. Havighurst and Janke found that children of higher family social status tended to do better in all of the tests than children of lower social position.

Hanson and Robinson (1967) sought to describe differences in reading readiness and reading achievement between three socioeconomic levels of children in kindergarten, first grade, second grade, and third grade. The disadvantaged subjects were enrolled in a public school in the Chicago area which is representative of a low socioeconomic strata. All the subjects in the sample were Black. The subjects in Hanson and Robinson's average group were Caucasian children from a middle-class school in a white suburb of Chicago. The advantaged subjects attended a private university-

affiliated school in Chicago. About 75 percent of the advantaged subjects were Caucasian; the remainder were Black or of Oriental descent.

Hanson and Robinson administered the Metropolitan Readiness Tests to the kindergarten and first graders during the first half of the school year. At the end of the school year, the Metropolitan Achievement Tests were administered in the first, second, and third grades.

Analysis of variance procedures were used to determine differences in reading readiness and reading achievement among the subjects of different socioeconomic levels at each grade level. At the kindergarten level, the advantaged subjects scored significantly higher than the average subjects, and the average subjects scored significantly higher than the disadvantaged on the Metropolitan Readiness Tests. The advantaged first graders also scored significantly higher than the average first graders on the Metropolitan Readiness Tests and the average first graders scored significantly higher than the disadvantaged first graders. The advantaged first graders scored significantly higher than the average first graders on the Metropolitan Achievement Tests and the advantaged first graders also scored significantly higher than the disadvantaged first graders on the Metropolitan Achievement Tests. The scores of the average first graders were significantly higher than the disadvantaged first graders on sections one and two of the Metropolitan Achievement Tests.

In both the second and third grades, Hanson and Robinson found that the advantaged subjects achieved significantly higher than the disadvantaged on the reading achievement tests. The scores between the advantaged group and the average group and the scores between the average group and the disadvantaged group were smaller and less uniform.

Hanson and Robinson concluded that reading readiness and reading achievement scores attained by the advantaged subjects are significantly higher in each grade than those attained by the disadvantaged. The authors concluded that their study contributes:

. . . additional evidence to show that advantaged youngsters enter school more ready to receive reading instruction and experience greater achievement in reading in the primary grades than other children.

In a similar study, Hill and Giammatteo (1963) sought to investigate socioeconomic status and its relationship to vocabulary achievement, reading comprehension, arithmetic skill, problem solving, and a composite of these variables. Their sample was 223 third-grade children from western Pennsylvania. The group was the total third grade population of nine classrooms and three schools, representing two school districts.

High and low socioeconomic status was determined by the use of an interview sheet completed by teachers and pupils. The Otis Quick-Scoring Mental Ability Test, Iowa Tests of Basic Skills, and the Scott-Foresman Basic Reading Tests were administered to all 223 third-grade children in the



study. Correlations were computed between socioeconomic status, reading comprehension, vocabulary, and arithmetic skills. These correlates were: 1. between socioeconomic status and reading comprehension,  $r = .90$ ; 2. between socioeconomic status and vocabulary,  $r = .77$ ; and 3. between socioeconomic status and arithmetic skills,  $r = .77$ .

The mean scores of the high and low socioeconomic status groups were compared for reading comprehension, vocabulary, and arithmetic skills by t-tests. The means obtained indicated the children from the high socioeconomic group by grade three were eight months ahead of the children from the low group in the area of vocabulary achievement. In reading comprehension achievement, the range between the groups was equivalent to a full school term or nine months. A full six months difference was evidenced between the high and low groups in arithmetic skills. These differences were significant.

Hill and Giammatteo concluded that the results of this study strengthen the accumulative evidence that socioeconomic status affects school achievement. All scholastic achievement areas reported in this study were effected by socioeconomic factors.

Vilscek and Cleland (1968) conducted a study for the purpose of examining the extended effects and outcomes of the Coordinated Basal Language Arts Approach and the Integrated Experience Approach to Communication on pupils' language development at the second and third grade levels.

One of the major purposes of the investigation was to determine the effects of socioeconomic levels on pupil achievement. The study began with 669 first graders in twenty-four first grades, selected from 241 first grades in the Pittsburgh Public Schools. This pupil population represented three distinct levels of socioeconomic strata as measured by the Hollingshead and Redlich Scale. Pupils and teachers were randomly assigned to classes within either teaching method. Testing at the end of the first grade reflected that in every instance pupils in the Integrated Experience Approach had significantly higher mean scores.

Eighteen of the original twenty-four first-grade classes constituted the pupil populations in second grade and third grade. There were 396 pupils in the analysis at the end of the second grade. There were 351 pupils in the analysis at the end of the third grade. The following instruments were administered for measurement: Stanford Achievement Tests - Advanced Primary Battery; the University of Pittsburgh Rating Record of Supervisory Attitudes and Pupils' Social Integration; the University of Pittsburgh Diagnostic Ratings of Language Outcomes for Second Graders, and Third Graders; Informally Devised Tests of Listening Comprehension Abilities; Minnesota Tests of Creative Thinking; the San Diego Pupils Inventory of Reading Attitudes; the Gilmore Oral Reading Test; the Gates Word Pronunciation Test; the Fry Phonetically Regular Words Oral Reading Test; the Restricted Stimulus Measure of Written Expression; and, Pupil

Inventories. A factorial analysis of variance was used to analyze the data. The analysis of reading achievement scores at the end of second grade yielded significant differences between pupils in the three socioeconomic levels. These same differences were also significant at the end of third grade. Vilscek and Cleland concluded that pupils' achievement was directly related to their socioeconomic level with higher socioeconomic status pupils having higher achievements than do lower socioeconomic status pupils.

In a longitudinal study to determine predictor variables of future success in reading and arithmetic, Hirst (1969) drew a sample of 300 kindergarten children from a population of beginning public school students to be tested in kindergarten, first, and second grade. These 300 children were selected from nine schools on the basis of socioeconomic status of the school. The schools represented lower socioeconomic, middle socioeconomic, and higher socioeconomic levels. All children in a kindergarten class of the selected schools were included. Reading achievement scores as well as many other variables commonly associated with academic achievement were analyzed using multiple regression analysis. Results of the research tend to indicate that socioeconomic status is a significant predictor variable of second grade reading and arithmetic success. Hirst concluded that socioeconomic status was a significant predictor of second grade reading success.

Oakland (1969) studied the relationship between reading achievement and social class membership as well as the relationship between reading achievement and performance on phonemic and non-phonemic auditory discrimination tests. Oakland's sample consisted of twenty randomly selected children in each of three socioeconomic status levels. The socioeconomic status of the children in these three groups, lower-lower class, upper-lower class, and upper-middle class was determined by ratings on the Index of Status Characteristics. The assessment of phonemic auditory discrimination was made with the Wepman Auditory Discrimination Test and non-phonemic auditory discrimination was measured using experimental tests. Reading achievement was determined by utilizing an Eight-Point Reading Scale which permitted teachers to specify the children's independent reading level on basal readers. Each subject was tested individually.

An analysis of covariance with I. Q. scores as the covariate was significant at the .05 level for differences in reading achievement between the different socioeconomic status groups. Children of the upper-middle class group read significantly better than children in the upper-lower class group, who in turn read significantly better than children in the lower-lower class group. Oakland concluded that the results of the study supported the widely held notion that there is a direct correspondence between social class status and reading achievement.

Although Bruininks (1970) was concerned with determining various characteristics of disadvantaged children other than reading achievement, he did examine the extent of differences between disadvantaged children and normative samples relative to reading achievement. Bruininks' sample consisted of 354 subjects, 177 boys and 177 girls. All 354 subjects were from low socioeconomic status homes and all subjects had been enrolled in the public schools of Metropolitan Nashville-Davidson County, Tennessee for three years.

In addition to measures of oral and written language and creative thinking abilities, each subject was administered the Metropolitan Achievement Tests at the end of the second and third grades. All testing was done by trained and/or qualified psychometrists in small group sessions. Bruininks found that on all MAT subtests, the subjects attained scores appreciably below those of the normative population. This disparity was of approximately 18 months. Bruininks interpreted these findings as indicating that disadvantaged children do not achieve as well as other children do in reading.

An examination of studies of reading achievement using socioeconomic status as an independent variable do conclude that reading achievement is related to socioeconomic status. In the studies reviewed, children from a higher socioeconomic status consistently achieved better reading scores when compared to children with lower socioeconomic status.

### Teacher Judgement of Reading Readiness

Henig (1949) conducted a study "to determine the comparative forecasting value of the Lee-Clark Reading-Readiness Test and of teachers' estimates of their pupils' likelihood of succeeding in learning to read." Ninety-eight first-grade children from a low socioeconomic area were ranked by their teacher on their relative likelihood to succeed in learning to read. A day or two later the Lee-Clark Reading-Readiness Test was administered to the 98 children by their teachers. A Chi Square value of 51.9 ( $p = .01$ ) led Henig to conclude that the forecasts made by teachers, versed in reading readiness techniques, have just as high a predictive value as does the Lee-Clark Reading-Readiness Test.

Kermoian (1962) utilized teacher ratings of reading readiness rather than rankings to determine the validity of teacher judgement of the readiness status of children entering first grade. The investigation was conducted in the fall of 1959. Participating in this study were 276 first-grade children from six elementary schools of differing socioeconomic areas of San Francisco. Thirteen first-grade teachers were used to obtain the ratings.

Although all cooperating teachers were familiar with the Metropolitan Readiness Tests, they were asked to carefully examine the manual and to study the sections dealing with a description of the subtests and of letter rankings. The teachers were asked to observe their children

for the first two weeks of school and at the end of the two week period were asked to rate each child according to the five point scale used on the Metropolitan Readiness Tests for Reading Readiness, Number Readiness, and Total Readiness. The Metropolitan Readiness Tests, Form R, were then administered.

The Pearson  $r$  was used to calculate the correlation coefficients. The correlations were .73 for Reading Readiness, .73 for Number Readiness, and .77 for Total Readiness. All three correlations were significant beyond the .01 level of confidence. Kermoian concluded that the classroom teacher's appraisal of readiness correlates highly and significantly with that of formal instrument and that most teachers can easily judge readiness status through non-test techniques.

In a study designed to determine how kindergarten teachers' evaluations of pupil reading readiness compared with reading readiness measures, Smith (1968) studied 360 pupils from eleven classrooms in the Longview, Washington School District. The kindergarten teachers were asked to assign reading readiness ratings to their pupils using the criteria and ratings of the Metropolitan Readiness Tests. Reading readiness classifications were then obtained using the Metropolitan Readiness Tests. The Metropolitan was teacher administered at the end of kindergarten. Reading achievement ratings were obtained from first-grade teachers during the second year of the study. The Gates-MacGinitie

Reading Tests were administered during May of the first grade. A Chi Square coefficient was calculated using a contingency table for each comparison. The Chi Square value comparing teacher ratings of reading readiness with the Metropolitan Readiness Tests' ratings were significant at the .01 level. Smith concluded that kindergarten teachers could provide a reading readiness evaluation that is apparently as accurate as commercial reading readiness tests.

Merrill (1968) had as his primary purpose, determining the relationship that exists between teachers' informal estimates of their pupils' reading readiness and first-grade reading achievement. A second purpose of Merrill's work was to determine the relationship between these informal estimates of reading readiness and a standardized reading readiness test.

Merrill's sample for the study consisted of 692 pupils in forty first-grade classrooms. The teachers were asked to observe their children for two weeks and to then make an informal estimate of their pupils' reading readiness. During the third week of school the Lee-Clark Reading Readiness Test was administered by the teachers. The Gates-MacGinitie Reading Test, Primary A, Form 1, was administered at the end of the year. This instrument was also teacher administered.

Pearson Product Moment Correlation coefficients were calculated between the teachers' informal estimates of reading readiness and first-grade reading achievement. The average correlation coefficient for the forty classrooms was



.537, which was significant at the .01 level. Although, Merrill found the average correlation to be significant, only twenty-eight of the correlations computed between the informal estimate of reading readiness and the Lee-Clark Reading Readiness Test scores were significant.

Merrill obtained similar results between the Lee-Clark Reading Readiness Test and first-grade reading achievement. The average correlation coefficient for the forty classrooms was .537, which was significant at the .01 level. Only twenty-eight of the individual classroom coefficients were significant. Merrill concluded that informal teacher ratings of reading readiness were as good as the standardized measure of reading readiness for his total sample but that these coefficients were not necessarily significant for individual classrooms.

Research supports the conclusion that teacher judgments of pupils' reading readiness status are as reliable an indicator of reading readiness as are standardized reading readiness test results. A change in teacher rankings of reading readiness as a result of receiving falsified reading readiness test scores would therefore be indicative of changed teacher expectations for success in reading.

#### Summary

The body of research leads one to conclude that the phenomenon of teacher expectations does exist and is an influential factor in determining how much is learned in the

primary school classroom. Although the research demonstrating teacher expectations has been primarily concerned with I. Q. gains, it has direct implications for reading instruction. If I. Q. scores can be increased by giving a teacher false expectations, it is a reasonable supposition to believe that other cognitive pupil gains can be achieved in the same manner.

An examination of studies of reading achievement using socioeconomic status as an independent variable do conclude that reading achievement is related to socioeconomic status. In the studies reviewed, children from a higher socioeconomic status consistently achieved better reading scores when compared to children with lower socioeconomic status.

Research supports the conclusion that teacher judgments of pupils' reading readiness status are as reliable an indicator of reading readiness as are standardized reading readiness test results. A change in teacher rankings of reading readiness as a result of receiving falsified reading readiness test scores would, therefore, be indicative of changed teacher expectations for success in reading.

## CHAPTER III

### DESIGN AND METHODOLOGY

#### Introduction

This chapter contains a description of the population of the study, the testing procedures, the instruments used in collecting the data, and the statistical treatment of the data.

#### Description of the Population

The population for this study consisted of 64 randomly selected first-grade pupils in the public schools of a north-eastern Oklahoma city who were of a high or low socioeconomic status and were either high or low reading readiness status pupils.

The experimental population were those children who met the following requirements:

1. Ranked in the top one-third or the bottom one-third of their class as rated on socioeconomic status by the Otis Dudley Duncan Socioeconomic Index for Occupations based on the NORC scale.

2. Received a reading readiness score of either one standard deviation (real limits) below the norm mean reading

readiness score on the Metropolitan Readiness Test, Form B, for that pupil's classroom.

3. Were not repeating the first-grade.

#### Testing Procedures

The classroom teachers first ranked the students in her classroom as to level of reading readiness. A form was provided for each classroom by this researcher. The teacher ranking of reading readiness was done during the third week of school. On the same form, the school secretary recorded the occupation for the head of the family. This was done after the teachers ranked reading readiness.

The Metropolitan Readiness Tests, Form B, was administered and scored by trained examiners from the Oklahoma State University Reading Center. In each testing situation only the examiner and the subjects were present.

Due to the failure of the intelligence test to arrive, it was decided to forgo the use of an intelligence test as a screening device. The results of the Metropolitan Readiness Tests were returned to the classroom teachers during the sixth week of school. The reported results included the scores of 32 children which had been falsified. These scores were falsified by adding an equal amount to each subtest score. The total added raised each of these 32 scores two standard deviations. One week after receiving the falsified test results, the teachers were asked to rank their students' reading readiness level for a second time.

During the first week in December, 1972, trained examiners from Oklahoma State University administered and scored the Stanford Achievement Test, Primary I Battery, Form W. The Stanford Achievement Test was administered to all the first-grade students in the 12 schools. The teachers were not present. The Short Test of Educational Abilities, Level I, was administered in February, 1973, by the classroom teachers and scored at the Oklahoma State University Reading Center to determine if the original sample was in the normal range of intelligence. Due to missing data on this instrument, (approximately 30%) the scores were used in no way.

Trained examiners from the Oklahoma State University Reading Center administered the first three subtests of the Stanford Achievement Test, Primary I, Form X, during the third week in April, 1973. Since only the first three subtests were to be used in the data analysis, the regular classroom teacher administered the remainder of the Stanford Achievement Test. The entire Stanford Achievement Test was scored by the examiners from Oklahoma State University.

#### Instruments Used in the Study

Short Test of Educational Ability,  
Level 1 (1966)

It was intended that this test be used to identify pupils who were within the average range of intelligence.

It measures three aspects of intelligence: cause and effect relations, problem solving, and spatial relations. The "What Would Happen If...?" subtest consists of twelve multiple choice items, each of which requires a selection from four pictorial alternatives to answer a question. The "How Would You...?" subtest consists of eight multiple choice items, each of which requires a selection from four pictorial alternatives to answer a question. The Spatial Relations subtest consists of twenty-five multiple choice items, in which the subject is required to choose the correct shape, from one of four shapes, necessary to complete the stimulus symbol (STEA, 1966).

Standardization was based on several thousand children by correlating with the Primary Mental Abilities. The Short Test of Educational Ability has a mean I. Q. of 100 and a standard deviation of 16 (STEA, 1966). In the present investigation children with STEA I. Q.'s of between 84 and 116, inclusive, were intended to be retained in the sample.

Metropolitan Readiness Tests,  
Form B (1966)

This test was devised to measure the extent to which school beginners have developed in the several skills and abilities that contribute to readiness for first-grade instruction. It is made up of six subtests which yield a total readiness score for the child. Primarily the test is designed to be administered at the end of the kindergarten

year or the first month of the first-grade year. The six subtests which make up this test are:

Test 1. Word Meaning, a 16-item picture vocabulary test. The pupil selects from three pictures the one that illustrates the word the examiner names.

Test 2. Listening, a 16-item test of ability to comprehend phrases and sentences instead of individual words. The pupil selects from three pictures the one which portrays a situation or event the examiner describes briefly.

Test 3. Matching, a 14-item test of visual perception involving the recognition of similarities. The pupil marks the one of three pictures which matches a given picture.

Test 4. Alphabet, a 16-item test of ability to recognize lower-case letters of the alphabet. The pupil chooses a named letter from among four alternatives.

Test 5. Numbers, a 26-item test of number knowledge.

Test 6. Copying, a 14-item test which measures a combination of visual perception and motor control (Hildreth, Criffits, McGauvran, 1966).

The normative population of the Metropolitan Readiness Tests included a nation-wide sample of approximately 15,000 pupils in 70 school systems. Norms for Form B of the Readiness Tests were established by means of a program equating results on this form with those of Form A for the same pupils (Hildreth et al., 1966).

The Metropolitan Readiness Tests, Form B, scores were used to determine the subject's level of reading readiness and as the manipulated independent variable in this study.

Stanford Achievement Tests: Primary I  
Battery, Form W and Form X (1964)

This test was used to measure pupil's reading achievement. It measures three aspects of reading achievement: word reading, paragraph meaning, and vocabulary. Only the reading subtests were analyzed in this study. The Word Reading Test consists of 35 multiple choice items, each of which requires a selection from four alternatives of the correct word which stands for a picture. The Paragraph Meaning Test contains 37 paragraphs (38 multiple choice items) of increasing difficulty from which one or more words have been omitted. The correct word for each omission must be selected from among four choices. The Vocabulary Test employs 39 multiple choice items in which the pupil is required to select from a series of three alternatives the proper answer to a question or a statement read by the teacher (Kelley, Madden, Gardner, & Rudman, 1964).

The tests were standardized on a minimum of 10,000 pupils per grade level. Content or curricular validity is based on the content of the typical elementary school curriculum. Typical courses of study and textbooks were examined by the authors; experimental tests were tried out and the items reviewed by a number of classroom teachers and reading specialists. Split-half reliability coefficients of .85 for the Word Reading Test, of .90 for the Paragraph



Meaning Test, and of .79 for the Vocabulary Test were reported for grade one (Kelley, et al., 1964).

National Opinion Research Center  
Scale (NORC)

Cecil North and Paul Hatt, in 1947, developed a method of occupational classification. The NORC measure of relative prestige for each occupation is a rank based on ordering the average scores derived from the weighted proportions of excellent, good, average, somewhat below average, or poor ratings for each occupation. The lowest occupation on this prestige scale is shoe shiner with a score of 33 while the highest score is 96 for United States Supreme Court Justice. Otis Dudley Duncan constructed a socioeconomic index from census information on detailed occupation characteristics. This index related the NORC to the 1950 Bureau of Census Report and allows a researcher to classify more varied occupations (Reiss & Duncan, 1961). The Otis Dudley Duncan Socioeconomic Index for Occupations based on the NORC was used in this study because it provides a system of classification that allowed the investigator to classify all occupations of the subjects' parents. The Otis Dudley Duncan Socioeconomic Index for Occupations was used to classify children as being of either high socioeconomic status or as being of low socioeconomic status.

## Statistical Design

The reading readiness scores and the Otis Dudley Duncan Socioeconomic Index for Occupations rankings were used to select the sample. The procedures for this are explained in the section on the Description of the Population.

A factorial analysis of variance with unequal cell frequencies will be used to determine the effects of socioeconomic status and reported reading readiness scores on the reading achievement (Winer, 1971).

The factorial arrangement is as follows:

B

A	$a_1b_1$	$a_1b_2$
	$a_2b_1$	$a_2b_2$

Where:

$a_1$  = high socioeconomic status

$a_2$  = low socioeconomic status

$b_1$  = experimental treatment (falsified scores)

$b_2$  = control (no falsified scores)

Hypotheses 3, 4, and 5 are concerned with the relationships that may exist between true reading readiness rankings, uninformed teacher rankings of reading readiness, and informed teacher ranking of reading readiness. An average



## CHAPTER IV

### TREATMENT OF DATA AND ANALYSIS OF RESULTS

The purpose of this study was to determine to what extent teacher expectations influence reading performance of students. Tests were made to determine if falsified reading readiness scores had any influence on reading achievement. The study included an analysis of the relationships that exist between teacher rankings of reading readiness and a ranking of actual reading readiness scores obtained on the Metropolitan Readiness Test. An analysis of teacher rankings of reading readiness before and after falsified Metropolitan Readiness Test scores were given to them was also done.

The hypothesis related to the effect falsified reading readiness scores have on reading achievement will be considered along with the hypothesis related to the interaction of falsified reading readiness and socioeconomic status. Next, the hypothesis concerning the relationship between teacher rankings of reading readiness and a ranking of Metropolitan Readiness Test scores will be examined. The hypotheses concerning the relationship that exists between teacher rankings of reading readiness after falsified reading

readiness test scores have been returned and the teachers' earlier rankings and the Metropolitan Readiness Test rankings will be examined last.

#### Tests of the Hypotheses

Hypothesis 1: There is no significant difference in reading achievement between groups with different levels of reading readiness who are perceived by the teacher as being at the same state of reading readiness performance.

Hypothesis 2: Student socioeconomic status and teacher knowledge of student reading readiness performance will not interact to shape reading achievement.

To test these hypotheses, and due to loss of subjects, a factorial analysis of variance for unequal cell frequencies was used. These two hypotheses were tested separately for the Word Reading subtest of the Stanford Achievement Test at each time the Stanford Achievement Test was administered. The Stanford Achievement Test was administered twice. These results are presented in Tables I-VI.

Tables I-VI reported the F values and the level of significance for each F value for the Word Reading subtest, the Paragraph Meaning subtest, and the Vocabulary subtest of the Stanford Achievement Test at each of two test administrations. An examination of Tables I-VI indicates that no F value reached the .05 level of significance necessary for a rejection of the null hypothesis. Therefore,

TABLE I

ANALYSIS OF VARIANCE OF DIFFERENCE BETWEEN WORD  
 READING FOR DIFFERENT SOCIOECONOMIC STATUS  
 AND FALSIFIED METROPOLITAN READINESS  
 TEST SCORES TWO MONTHS AFTER DATA  
 PRESENTED TO TEACHERS

Source	SS	df	MS	F
Socioeconomic Status	9.779	1	9.779	
Reported Readiness	18.142	1	18.142	1.945**
Interaction	12.077	1	12.077	1.295*
Within Cells Error	447.568	48	9.324	
Total	487.566	51		

\*ns  
 \*\*p = .20

TABLE II

ANALYSIS OF VARIANCE OF DIFFERENCE BETWEEN PARAGRAPH  
 MEANING FOR DIFFERENT SOCIOECONOMIC STATUS AND  
 FALSIFIED METROPOLITAN READINESS TEST SCORES  
 TWO MONTHS AFTER DATA PRESENTED TO TEACHERS

Source	SS	df	MS	F
Socioeconomic Status	23.663	1	23.663	
Reported Readiness	18.517	1	18.517	1.395*
Interaction	4.071	1	4.071	0.306*
Within Cells Error	636.749	48	13.265	
Total	683.000	51		

\*ns  
 \*\*p = .20

TABLE III

ANALYSIS OF VARIANCE OF DIFFERENCE BETWEEN VOCABULARY  
FOR DIFFERENT SOCIOECONOMIC STATUS AND FALSIFIED  
METROPOLITAN READINESS TESTS SCORES TWO MONTHS  
AFTER DATA PRESENTED TO TEACHERS

Source	SS	df	MS	F
Socioeconomic Status	189.789	1	189.789	
Reported Readiness	2.575	1	2.575	0.078*
Interaction	8.479	1	8.479	0.257*
Within Cells Error	1583.215	48	32.983	
Total	359.068	51		

\*ns

TABLE IV

ANALYSIS OF VARIANCE OF DIFFERENCE BETWEEN WORD  
READING FOR DIFFERENT SOCIOECONOMIC STATUS  
AND FALSIFIED METROPOLITAN READINESS  
TESTS SCORES SIX MONTHS AFTER  
DATA PRESENTED TO TEACHERS

Source	SS	df	MS	F
Socioeconomic Status	147.019	1	147.019	
Reported Readiness	90.653	1	90.653	2.320**
Interaction	0.367	1	0.122	0.003*
Within Cells Error	1718.975	44	39.067	
Total	1957.014	47		

\*ns

\*\*p = .20

TABLE V

ANALYSIS OF VARIANCE OF DIFFERENCE BETWEEN PARAGRAPH  
MEANING FOR DIFFERENT SOCIOECONOMIC STATUS AND  
FALSIFIED METROPOLITAN READINESS TESTS SCORES  
SIX MONTHS AFTER DATA PRESENTED TO TEACHERS

Source	SS	df	MS	F
Socioeconomic Status	38.275	1	38.275	
Reported Readiness	11.976	1	11.976	0.308*
Interaction	0.303	1	0.303	0.002*
Within Cells Error	1706.997	44	38.795	
Total	1757.551	47		

\*ns

TABLE VI

ANALYSIS OF VARIANCE OF DIFFERENCE BETWEEN VOCABULARY  
FOR DIFFERENT SOCIOECONOMIC STATUS AND FALSIFIED  
METROPOLITAN READINESS TESTS SCORES SIX MONTHS  
AFTER DATA PRESENTED TO TEACHERS

Source	SS	df	MS	F
Socioeconomic Status	52.389	1	52.389	
Reported Readiness	30.723	1	30.723	1.500*
Interaction	60.171	1	60.171	2.941**
Within Cells Error	900.170	44	20.458	
Total	1043.453	47		

\*ns

\*\*p = .10



no components of Hypothesis 1 or Hypothesis 2 can be rejected by the findings of this study.

Hypothesis 3: There is no significant relationship between a teacher ranking of reading readiness status of her children before readiness tests have been administered and a ranking of the children's Metropolitan Readiness Test scores.

To test this hypothesis a Spearman Rank-Order correlation coefficient was computed between the ranking of reading readiness prior to the administration of the Metropolitan Readiness Tests and a ranking of actual reading readiness scores. A correlation coefficient was calculated for each classroom. The average Spearman Rank-Order correlation coefficient was then calculated for the 26 classrooms (Guilford, 1965). The data is presented in Table I. The average Spearman Rank-Order coefficient for the 26 classrooms is 0.78 and has a critical value exceeding the .01 level (Segal, 1956).

Table VII reported the average Spearman Rank-Order correlation coefficient between the Metropolitan Readiness Tests, Form B, and rankings of reading readiness by the classroom teacher prior to the administration of the Metropolitan Readiness Tests. An inspection of Table VII indicates that the average correlation between the rankings of the Metropolitan and the teacher rankings of reading readiness was .78. This correlation is significant beyond the .01 level of confidence. Thus, the null hypothesis of

TABLE VII

SPEARMAN RANK ORDER CORRELATIONS OF METROPOLITAN  
 READINESS TEST RANKINGS AND PRETEST TEACHER  
 RANKINGS OF READING READINESS

Classroom	N	r	t	p	z
A	22	0.747	5.024	0.001	0.97
B	22	0.902	9.327	0.001	1.47
C	23	0.959	15.440	0.001	1.95
D	22	0.657	3.894	0.001	0.79
E	27	0.585	3.609	0.01	0.68
F	18	0.832	5.996	0.001	1.19
G	19	0.606	3.140	0.01	0.71
H	20	0.789	5.444	0.001	1.07
I	19	0.877	7.522	0.001	1.38
J	16	0.380	1.536	0.5	0.40
K	17	0.821	5.568	0.001	1.16
L	15	0.679	3.334	0.01	0.83
M	16	0.803	5.034	0.001	1.10
N	19	0.789	5.291	0.001	1.07
O	20	0.919	9.870	0.001	1.59
P	20	0.696	4.110	0.001	0.87
Q	21	0.818	6.197	0.001	1.16
R	22	0.574	3.133	0.01	0.65
S	19	0.759	4.775	0.001	1.00
T	15	0.646	3.050	0.01	0.78
U	25	0.819	6.837	0.001	1.16
V	23	0.897	9.283	0.001	1.42
W	23	0.910	10.054	0.001	1.53
X	23	0.419	2.113	0.05	0.45
Y	24	0.781	5.857	0.001	1.05
Z	26	0.656	4.256	0.001	0.79
Total	536				27.22
				Average z =	1.05
				Average r =	.78*

\*p = .001

no significant relationship between a teacher ranking of reading readiness status of her students before readiness tests have been administered and a ranking of the student's Metropolitan Readiness Test scores can be rejected.

Hypothesis 4: There is no significant relationship between a teacher ranking of reading readiness status of her students before readiness tests have been administered and teacher ranking of reading readiness status after reading readiness tests have been administered.

Hypothesis 5: There is no significant relationship between a teacher ranking of reading readiness status of her students after readiness tests have been administered and a ranking of the children's Metropolitan Readiness Test scores.

To test these hypotheses, a second teacher ranking of each teacher's student's reading readiness after the teacher had been given falsified reading readiness test results was necessary. This ranking was done. However, due to the movement of children from classroom to classroom the second rankings constituted a slightly different sample for all but three teachers. Therefore, average correlation coefficients to test Hypothesis 4 and Hypothesis 5 were calculated using the rankings of these three teachers. These results are presented in Table VIII and Table IX.

The observed relationship for Hypothesis 4 is presented in Table VIII. Using  $p = .05$  as the recognized level for rejecting the null hypothesis, Hypothesis 4 can be rejected on the basis of the evidence presented.

TABLE VIII

SPEARMAN RANK ORDER CORRELATIONS OF PRETEST TEACHER  
RANKING OF READING READINESS AND POSTTEST TEACHER  
RANKINGS OF READING READINESS

Classroom	n	r	t	p	z
G	19	.799	2.630	.02	1.10
L	15	.115	.417	ns	.12
Y	24	.103	.485	ns	$\frac{.11}{1.33}$
				Average z = .44	
				Average r = .41*	

\*p = .01

TABLE IX

SPEARMAN RANK ORDER CORRELATIONS OF METROPOLITAN  
READINESS TEST RANKINGS AND POSTTEST TEACHER  
RANKINGS OF READING READINESS

Classroom	n	r	t	p	z
G	19	.822	5.944	.001	1.16
L	15	.543	2.330	.05	.61
Y	24	.789	6.018	.001	$\frac{1.07}{2.84}$
				Average z = .95	
				Average r = .74*	

\*p = .01

The observed relationship for Hypothesis 5 is presented in Table IX. Using  $p = .05$  as the recognized level for rejecting the null hypothesis, Hypothesis 5 can be rejected on the basis of the evidence presented.

### Summary

This chapter has presented a detailed analysis of the statistical treatment of the data. The following hypotheses were rejected:

1. There is no significant relationship between a teacher ranking of reading readiness status of her children before readiness tests have been administered and a ranking of the children's Metropolitan Readiness Test scores.

2. There is no significant relationship between a teacher ranking of reading readiness status of her students before readiness tests have been administered and teacher ranking of reading readiness status after reading readiness tests have been administered.

3. There is no significant relationship between a teacher ranking of reading readiness status of her students after readiness tests have been administered and a ranking of the children's Metropolitan Readiness Test scores.

The following hypotheses could not be rejected:

1. There is no significant difference in reading achievement between groups with different levels of reading readiness who are perceived by the teacher as being at the same state of reading readiness performance.

2. Student socioeconomic status and teacher knowledge of student reading readiness performance will not interact to shape reading achievement.

## CHAPTER V

### SUMMARY AND CONCLUSIONS

#### General Summary of the Investigation

This study was concerned with determining if falsified reading readiness test scores significantly changed reading achievement. A second concern of this investigation was to determine the relationship that exists between teacher rankings of reading readiness and the Metropolitan Readiness Tests, Form B, and attempted to determine if teacher rankings of reading readiness are influenced by reading readiness test results.

The sample consisted of 64 randomly selected first-grade children from 26 classrooms who ranked one standard deviation (real limits) below the mean Metropolitan Readiness Tests, Form B, score for individual's classroom and who ranked in either the top or bottom third of their classroom when ranked by the Otis Dudley Duncan Socioeconomic Index for Occupations. No children repeating the first grade were included in the sample.

The Stanford Achievement Tests, Primary I Battery, has six subtests. The Stanford was administered twice. Form W of the Stanford was administered approximately two months

after reading readiness test results were given to the classroom teachers. Form X was administered six months after reading readiness test results were given to the classroom teacher. A factorial analysis of variance for unequal cell frequencies was run of the data from the first three Stanford subtests each time the Stanford was administered.

Teacher rankings of reading readiness were obtained from the teachers before and after a reading readiness test was administered. A ranking of reading readiness was ascertained from the total score obtained on a standardized readiness test, the Metropolitan Readiness Tests, Form B. A ranking of socioeconomic status for each classroom was made from information provided by the schools utilizing the Otis Dudley Duncan Socioeconomic Index for Occupations based on the North-Hatt Scale (NORC). All six subtests of the Stanford Achievement Test, Primary I Battery, Form W and Form X, were administered but only the first three subtests of each battery were analyzed.

An average correlation coefficient was calculated between the Metropolitan ranking of reading readiness and the reading readiness ranking made by teachers prior to the administration of the readiness test. The other planned comparisons were not made by reason of not being able to meet statistical assumptions.



## Conclusions

When the results of the factorial analysis of variance are examined, one has to conclude that the lack of significance at an acceptable level for the rejection of hypotheses causes this study to fail to support the efficacy of the concept of the self-fulfilling prophecy. This seems to indicate that the effect of teacher expectations may be more subtle and their effect may not be as immediate as other studies have seemed to indicate.

Results of the study indicate that teacher rankings of reading readiness prior to the administration of a readiness test correlate significantly with the Metropolitan Readiness Tests, Form B. For the sample of teachers considered in this study, it can be concluded that the teacher's rankings of reading readiness are as reliable an indicator of the students' readiness status as are the Metropolitan Readiness Tests, Form B.

This study failed to indicate that teacher expectations for a child's probable success in reading can be demonstrated by altering her views of the child's reading readiness by the presentation of falsified readiness data. The significant average correlation coefficient between the two teacher rankings of pupil readiness would indicate that the teachers did not significantly change their rankings of their children's readiness as a result of receiving falsified readiness scores. One reason for the failure of this

study to demonstrate a measurable change in teachers' expectations may be a result of the small number of subjects in each classroom for whom false readiness scores were reported. In the three classrooms from which usable data was obtained for measuring the phenomenon of changing teacher expectations, a total of only four children's readiness scores were falsified. This could then only result in minor changes in the second teacher ranking and the statistical techniques available are not sensitive to such minute changes in rankings.

A second explanation for the lack of a demonstrable change in teacher expectations may have been due to delays in initiating the study. Because of unexpected problems at the beginning of the school year, the initiation of this study was delayed several weeks. The teachers may, therefore, have already formed their expectations of each child's probabilities for successfully learning to read before this researcher could provide them with data intended to bias their beliefs for the children's success.

A third factor for failing to demonstrate a change in teacher expectations through a change in teacher rankings of readiness may have been due to the teachers lack of experience with reading readiness tests. Reading readiness tests had never been given to all first-grade children in this school system previous to this study. Therefore, the teachers may not have placed much value in their predictive validity for success at learning to read, rather they may

be better attuned to other factors in making their judgments about children's reading readiness and did not rely on the data given them.

#### Recommendations

1. It is recommended that a study be made that would attempt to measure if teacher expectations are changed due to the presentation of falsified reading readiness scores. Falsified scores should be presented in such number per classroom as to better ensure a probable change in teacher rankings that would be statistically measurable. Any future study that attempts to induce changes in teacher expectations should be initiated as early in the school year as is possible. Creditability in the independent variables should be strived for by the researcher.

2. It is recommended that this study be replicated with the collection of data over several years to measure the long range effects of falsified reading readiness scores. The influence of falsified reading readiness scores may be more subtle than previously believed and their influence may be more pronounced in later grades.

## BIBLIOGRAPHY

- Akers, J. C. A Predictive Validity Study of the Metropolitan Readiness Tests. Unpublished doctoral dissertation, Oklahoma State University, 1967.
- Association for Supervision and Curriculum Development (ASCD). Perceiving, Behaving, Becoming. Washington, D. C.: NEA, 1962.
- Beez, W. V. Influence of Biased Psychological Reports on Teacher Behavior and Pupil Performance. Proceedings of the 76th Annual Convention of the American Psychological Association, 1968, 605-606.
- Bond, G. L., & Tinker, M. A. Reading Difficulties: Their Diagnosis and Correction. (2nd ed.) New York: Appleton-Century-Crofts, 1967.
- Brophy, J. E., & Good, T. L. Teacher's Communication of Differential Expectation for Children's Classroom Performance: Some Behavioral Data. Journal of Educational Psychology, 1970, 61, 365-371.
- Bruininks, R. H. Measures of Intelligence, Language, Creativity, Reading, and Writing Language Achievement of Disadvantaged Children. In J. A. Figurel (Ed.), Reading Goals for the Disadvantaged. Newark, Delaware: International Reading Association, 1970.
- Bruning, J. L., & Kintz, B. L. Computational Handbook of Statistics. Atlanta: Scott, Foresman, 1968.
- Claiborn, W. L. Expectancy Effects in the Classroom: A Failure to Replicate. Journal of Educational Psychology, 1969, 60, 377-383.
- Conn, L. K., Edwards, C. N., Rosenthal, R., & Crowne, D. Perception of Emotion and Response to Teacher's Expectancy by Elementary School Children. Psychological Reports, 1968, 22, 27-34.
- Erikson, E. H. Identity: Youth and Crisis. New York: Norton, 1968.

- Goldsmith, J. S., & Fry, E. The Effect of a High Expectancy Prediction on Reading Achievement and I. Q. of Students in Grade 10. Paper presented at the meeting of the American Educational Research Association, New York, N. Y., February, 1971.
- Good, T. L. Which Pupils Do Teachers Call On? Elementary School Journal, 1970, 70, 190-198.
- Guilford, J. P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill, 1965.
- Hanson, E., & Robinson, H. A. Reading Readiness and Achievement of Primary Grade Children of Different Socio-economic Strata. The Reading Teacher, 1967, 21, 52-56, 79.
- Havighurst, R. J., & Janke, L. L. Relations Between Ability and Social Status in a Midwestern Community. I: Ten-Year-Old Children. Journal of Educational Psychology, 1944, 35, 357-368.
- Henig, M. S. Predictive Value of a Reading Readiness Test and of Teachers' Forecasts. Elementary School Journal, 1949, 50, 41-46.
- Hildreth, G. H., Griffiths, N. L., & McGauvran, M. E. Metropolitan Readiness Tests. New York: Harcourt, Brace & World, 1966.
- Hill, E. H., & Giammatteo, M. C. Socio-Economic Status and Its Relationship to School Achievement in the Elementary School. Elementary English, 1963, 40, 265-270.
- Hirst, W. E. Identification in the Kindergarten of Factors That Make for Future Success in Reading and Identification and Diagnosis in the Kindergarten of Potential Reading Disability Cases. Final Report. ERIC Document. Ed 029710, 1969.
- Jose, J. Teacher-Pupil Interaction as it Relates to Attempted Changes in Teacher Expectancy of Academic Ability and Achievement. Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, Minnesota, March, 1970.
- Karlan, R. The Prediction of Reading Success and Reading Readiness Tests. Elementary English, 1957, 34, 320-322.
- Kelley, R. L., Madden, R., Gardner, E. F., & Rudman, H. C. Stanford Achievement Tests. New York: Harcourt, Brace & World, 1964.

- Kermoian, S. B. Teacher Appraisal of First-Grade Readiness. Elementary English, 1962, 39, 196-201.
- Merrill, J. D. An Investigation of Teachers' Informal Estimates of First-Grade Reading Readiness. Unpublished doctoral dissertation, University of Oregon, 1968.
- Oakland, T. D. Auditory Discrimination and Socioeconomic Status as Correlates of Reading Ability. Journal of Learning Disabilities, 1969, 2, 324-329.
- Palardy, J. M. The Effect of Teachers' Beliefs on the Achievement in Reading of First-Grade Boys. Unpublished doctoral dissertation, Ohio State University, 1968.
- Reiss, A. J., Jr., & Duncan, O. D. Occupations and Social Class. New York: Crowell-Collier, 1961.
- Rosenthal, R., & Jacobson, L. Teachers' Expectancies: Determinants of Pupils' I. Q. Gains. Psychological Reports, 1966, 19, 115-118.
- Schrank, W. R. The Labeling Effect of Ability Grouping. Journal of Educational Research, 1968, 62, 51-52.
- Schrank, W. R. A Further Study of the Labeling Effect of Ability Grouping. Journal of Educational Research, 1970, 63, 358-360.
- Short Test of Educational Abilities. Chicago: Science Research Associates, 1966.
- Siegel, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.
- Simpson, D. Perceptual Readiness and Beginning Readiness. Unpublished doctoral dissertation, Purdue University, 1960.
- Smith, M. H. Kindergarten Teachers' Judgements of Their Pupils' Readiness for Reading Instruction Compared with Readiness Test Results and First-Grade Achievement Measures. Unpublished doctoral dissertation, University of Oregon, 1968.
- Stauffer, R. G. Some Tidy Generalizations. In R. G. Stauffer (Ed.), The First Grade Reading Studies: Findings of Individual Investigations. Newark, New Jersey: International Reading Association, 1967.
- Vilscek, E. C., & Cleland, D. L. Two Approaches to Reading Instruction. Final Report. ERIC Document. Ed 022647, 1968.

Winer, B. J. Statistical Principles in Experimental Design.  
(2nd. Ed.) New York: McGraw-Hill, 1971.

## VITA

David V. Elijah, Jr.

Candidate for the Degree of

Doctor of Education

Thesis: TEACHER EXPECTATIONS: DETERMINANTS OF PUPILS'  
READING ACHIEVEMENT

Major Field: Elementary Education

Biographical:

Personal Data: Born at Rensselaer, Indiana, April 27,  
1942, the son of D. Victor and Lois M. Elijah.

Education: Attended elementary school in Ade, Indiana;  
attended secondary school in Morocco, Indiana;  
graduated from Morocco High School in 1960; re-  
ceived the Bachelor of Arts degree from Purdue  
University in 1966 with a major in Elementary Edu-  
cation; received the Master of Science degree from  
Oklahoma State University in 1972 with a major in  
Elementary Education; completed requirements for  
the Doctor of Education degree in May, 1974.

Professional Experience: Taught fifth grade arithmetic  
at Central Elementary School, Belen, New Mexico,  
1966-1967; fourth grade teacher and first grade  
teacher at Emerson Elementary School, Albuquerque,  
New Mexico, 1967-1970; remedial reading teacher at  
Jemez Springs School, Jemez Springs, New Mexico,  
summer, 1969; remedial reading teacher at Cushing  
Public Schools, Cushing, Oklahoma, 1970-1971;  
served as a graduate assistant at Oklahoma State  
University, 1971-1974; instructor of College Read-  
ing Improvement Classes, Oklahoma State University,  
1971; instructor of High School Reading Improve-  
ment Classes, Drumright, Oklahoma, 1971; Super-  
visor, Reading Clinic, Oklahoma State University,  
1972-1974; Supervisor, Reading Practicum, Oklahoma  
State University, 1972; Adjunct professor in Edu-  
cational Extension, spring, 1972, and fall, 1972;  
Assistant Project Director for Research Foundation  
Grants, Oklahoma State University, 1972-1974;



instructor of Developmental Reading in the Elementary School, Oklahoma State University, 1972; instructor of Reading Diagnosis, Oklahoma State University, 1972 and 1973; Elementary Consultant, Cleo-Aline Consolidated Schools, Cleo Springs, Oklahoma, 1973; Consultant to Special Services, evaluated the Oral Language Component to the Bilingual Teacher Education Program, New Mexico Highlands University, Las Vegas, New Mexico, 1973; Director of the Reading Center and Visiting Professor of Education, Eastern New Mexico University, Portales, New Mexico, summer, 1973; instructor of Nature and Characteristics of the Mentally Retarded Child, Oklahoma State University, 1973; Consultant and Diagnostician to the Tulsa County Early Childhood Education Project, Title III, ESEA, Tulsa, Oklahoma, 1973-1974.