This dissertation has been
microfilmed exactly as received

69-17,653

KESTER, Scott Woodrow, 1937-
THE COMMUNICATION OF TEACHER EXPECTATIONS AND THEIR EFFECTS ON THE ACHIEVEMENT AND ATTITUDES OF SECONDARY SCHOOL PUPILS.

The University of Oklahoma, Ph.D., 1969
Education, psychology

University Microfilms, Inc., Ann Arbor, Michigan

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/2570

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE COMMUNICATION OF TEACHER EXPECTATIONS AND THEIR EFFECTS ON THE ACHIEVEMENT AND ATTITUDES OF SECONDARY SCHOOL PUPILS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

SCOTT WOODROW KESTER

Norman, Oklahoma

1969

THE COMMUNICATION OF TEACHER EXPECTATIONS AND THEIR EFFECTS ON THE ACHIEVEMENT AND ATTITUDES OF SECONDARY SCHOOL PUPILS

APPROVED BY

Henry Angelino

P. J. Rupiper

Robert A. Smouse

Glenn R. Snider

DISSERTATION COMMITTEE

PLEASE NOTE: Appendix pages
are not original copy. Print
is indistinct on many pages.
Filmed in the best possible
way.

UNIVERSITY MICROFILMS.

ACKNOWLEDGMENTS

As the chairman of my dissertation committee, Dr. Henry R. Angelino has given both his time and encouragement. I thank him. I am also deeply grateful to Dr. Omer J. Rupiper for his interest and assistance, especially in the experimental design and analysis of the data; to Dr. Glenn R. Snider for his constant encouragement during my course work as well as my dissertation; and to Dr. Albert D. Smouse for his interest and assistance.

I would like to extend a special thanks to Dr. George A. Letchworth, my friend and colleague, for his encouragement, inspiration, and assistance throughout my program. I also thank his wife, Carol, for her interest and encouragement, especially during my language proficiency program.

To those individuals who helped me gather my data--Robert Royer, Stanley Bryant, June Matlock, Charles Matlock, Edward Porter, and Maxine Dunlop--I give special thanks. I also wish to thank my student assistant, Patricia Pruitt, for her help in recording the data. The typing of the dissertation was done by Kay Tims, and to her I give a very special thanks.

All tributes would be incomplete without acknowledging

the seventh grade pupils, the teachers, the counselors, and the administrators at Carl Albert, Jarman, and Monroney Junior High Schools in Midwest City, Oklahoma; Hefner and Western Oaks Junior High Schools in Bethany, Oklahoma; and Central Junior High School in Norman, Oklahoma. Their co-operation is deeply appreciated.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	iii
TABLE OF CONTENTS	v
LIST OF TABLES	vi
LIST OF ILLUSTRATIONS	vii
Chapter	
I. INTRODUCTION	1
II. METHOD	19
III. RESULTS	32
IV. DISCUSSION	45
V. SUMMARY	59
REFERENCES	64
APPENDIX A	71
APPENDIX B	73
APPENDIX C	74
APPENDIX D	75
APPENDIX E	76
APPENDIX F	81
APPENDIX G	86
APPENDIX H	87
APPENDIX I	90

LIST OF TABLES

Table		Page
1.	Analysis of Variance Performed on the Time Teachers Spent Communicating in a Positive-Accepting-Supportive Manner	35
2.	Analysis of Variance Performed on the Occurrences of Teachers' Communicating in a Positive-Accepting-Supportive Manner	36
3.	Analysis of Variance Performed on the Amount of Time Teachers Spent Communicating with Experimental Pupils vs. Control Pupils	38
4.	Linear Trend Analysis Performed on the Relationship Between Positive P-T Communication and Positive T-P Communication	38
5.	Analysis of Variance Performed on the Gain in Pupil Achievement in Language Arts as Opposed to Mathematics	39
6.	Analysis of Variance Performed on Pupil Attitudes Toward Teacher, Self, School, Mathematics (M), English (LA), and Fellow Students	41
7.	D Matrix for Teacher Attitudes	42

LIST OF ILLUSTRATIONS

Figure		Page
1.	Categories of the Category System of Observation	25
2.	Histogram of the Teachers' Knowledge of the Experiment	44

THE COMMUNICATION OF TEACHER EXPECTATIONS AND THEIR EFFECTS ON THE ACHIEVEMENT AND ATTITUDES OF SECONDARY SCHOOL PUPILS

CHAPTER I

INTRODUCTION AND PROBLEM

Since Merton's (1948) article on the "self-fulfilling prophecy," there has been an accumulation of data which seems to substantiate its existence. Self-fulfilling prophecy is generally defined by most authors as one person's expectations about the behavior of another person actually being fulfilled by that other person. Literature regarding expectancy behavior in psychology and education is scant; however, recent research evidence presented by Rosenthal and Jacobson (1968) indicates that some of our educational practices (e. g., grouping, tracking, disseminating test scores) should be re-evaluated in the light of expectancy behavior.

Review of the Literature

The occurrence of the self-fulfilling prophecy. The theoretical basis for the construct of expectancy behavior was established by Professor William I. Thomas in his theo-

rem of the situational limitations of behavior: "If men define situations as real they are real in their consequences" (1951, p. 81). Robert K. Merton, a sociologist, was one of the first behavioral scientists to capitalize on the implications of the construct. Merton (1948) described the self-fulfilling prophecy as a phenomenon which, in the beginning, is a "false" definition of a situation and which elicits a new behavior that, in turn, makes the originally false definition come "true." Rosenthal and Jacobson (1968) defined such an occurrence as "one person's expectation for another person's behavior unwittingly becoming a more accurate prediction simply for its having been made" (p. vii).

The effects of expectations of the effects of one person on the behavior of another have been reported in several areas not directly related to behavioral science. Whyte (1943) reported the phenomenon in the bowling behavior of the members of a street-corner gang. In a study in a large industrial plant, the performance of employees on tasks of dexterity favored their foremen's expectations (Bavelas, 1965). Beecher (1966), Shapiro (1960), and Sheard (1963) have reported placebo effects in the field of medicine. Finally, Allport (1950) has applied the concept to international affairs by suggesting that one nation's preparing for war begins the self-fulfilling prophecy cycle.

Efforts at identifying expectancy effects in the behavioral sciences have been relatively recent and the litera-

ture is limited. In survey research Harvey (1938) found that when interviewers were given fictitious information about their subjects, the subjects were assessed in accordance with the interviewers' expectations. Likewise, Wyatt and Campbell (1950) had some 200 interviewers prophesy the percentages of responses which would occur in answer to five questions on a survey. The interviewers tended to get responses in the proportions which they had prophesied.

Expectancy behavior has also received little attention from experimental psychology. Levy and Orr (1959) conducted studies to establish the validity of the Rorschach ink-blot technique of personality assessment. The results revealed that the experimenters who were more interested in construct validity obtained results which were more favorable to construct validation, whereas those who were interested in criterion validity achieved results favorable to their interests. Stanton and Baker (1942) conducted an experiment in which experimenters were testing the retention of geometric figures by 200 students. Each experimenter was given a key of "correct" responses (different responses were correct on different keys). Subjects responded in accordance with the "correct" responses marked on the keys. The findings of Stanton and Baker failed to be replicated, however, in studies conducted by Friedman (1942) and Lindzey (1951).

The majority of the research and writing in experimenter effects and expectancy behavior has been done by Robert Ro-

senthal and his associates. In an early study Rosenthal and Halas (1962), reported that experimenter effects were present in their research with invertebrates. Rosenthal and Fode (1963) found evidence of experimenter bias in a study using albino rats as subjects. In a later effort, Rosenthal and Lawson (1964), experimenter bias in a laboratory experiment with rats was again isolated. However, Rosenthal's (1966) more recent research efforts are of more immediate concern to the present study. One of the first studies conducted by Rosenthal (1966) was an experiment in which subjects were asked to rate the degree of success or failure of people shown in a series of ten photographs. Half the experimenters were told that people generally rated the photos as unsuccessful. Experimenters tended to obtain the results that they expected to obtain. Subsequent experiments were designed so as to learn something about the conditions which modify the effects of expectancy behavior (Rosenthal & Jacobson, 1968).

In an attempt to relate intellectual performance and learning to experimenter expectancies, Marwit and Marcia (1967) designed a study to investigate whether the number of responses given by a subject to a series of ink-blots was a function of the examiner's expectation or the subject's intellect. The results showed that those examiners prophesying greater response productivity obtained significantly more responses that did those prophesying fewer responses.

Masling (1965) reported that experimenter expectations also significantly influence the type of responses to ink-blots.

Larrabee and Kleinsasser (1967) investigated the effects of an examiner's expectancies on subjects taking a standardized test of intelligence. One examiner was told that the subjects were of above-average intelligence, and another examiner was told that the subjects were of below-average intelligence. The influence of examiner expectancy was significantly biased in the direction of the instructions.

Disadvantaged children and school performance. Several investigators have identified educational self-fulfilling prophecy as a significant variable in the low achievement of pupils functioning under conditions of socio-economical, cultural and/or educational deprivation. "Disadvantaged children by definition come from lower socio-economic groups where low income is married to values alien to the school culture" (Rosenthal & Jacobson, 1968, p. 48).

Sexton (1961) studied the relationship between income and educational opportunity and found that, in families whose average income was above $7000, achievement exceeded grade level; whereas, if the income was below $7000, achievement was below grade level. Havighurst (1965) also found a direct relationship between the level of family income and the level of academic achievement of the children in the home.

Another variable which influences the achievement of the disadvantaged child is that of achievement motivation.

McClelland (1961) pointed out that there are cultural and class differences in family commitment to achievement. However, Deutsch (1963) suggested that it is in the school where negative motivation for achievement is most inculcated. Passow (1963) reported teachers in lower-class schools did not set as high standards as did the teachers in the middle-class schools, and they also did not attempt to bring their pupils up to standard grade level.

Another explanation for the academic non-achievement among the disadvantaged is that they are restricted in language and cultural experiences which seem to enhance one's ability to learn in school. Deutsch (1963) found that lower-class children have not learned to pay attention, to adequately discriminate the sounds and sights around them. Loban (1964) found a consistent relationship between social class and ability to communicate.

In one report [Harlem Youth Opportunities Unlimited, Inc., (HARYOU), 1964], it was suggested that minority ethnic groups are especially likely to suffer from unfavorable halo effects in the teacher's evaluation. Cahen (1966), in an experiment in which he tested whether false information about pupil's aptitudes would influence the teacher's scoring of the pupil's test papers, found that the allegedly brighter pupils were given the benefit of the doubt to a much greater degree than the allegedly duller pupils.

Teacher expectations and pupil response. The teacher

acquires certain expectations about the performance of pupils from a variety of sources: the apparent socio-economic level of the pupil; the pupil's skin color; or information, whether true or false, about his previous performance. One of the most important sources of teachers' expectations about their pupils' academic abilities comes from their knowledge of pupil scores on various standardized tests (Deutsch et al., 1964; Gibson, 1965). The validity of a single test score of ability or achievement is well-known to be questionable; yet, no doubt, thousands of educators make decisions of considerable consequence to the pupil based on test scores.

Pitt (1956) proposed that scores on ability tests, when communicated to the teacher, affected the teacher's expectations regarding the pupils' performance, and that those expectations may then become self-fulfilling prophecies. Pitt's sample consisted of 165 fifth-grade boys with average or above-average IQ as indicated on a standardized test of intelligence. IQ's for one-third of the sample were reported to the teachers accurately; one-third were raised ten points; and one-third were lowered ten points. At the end of the school year the three groups were compared on school grades, achievement test results, teacher ratings and self-ratings. Pitt found essentially no effects on the results of objective achievement tests of the arbitrarily raised or lowered IQ's. It must be noted, however, that Pitt's teachers had taught their pupils for seven or eight weeks before being given the

expectancy communications (IQ scores). However, Pitt did find that there were differences in the pupil's attitudes about themselves, their school work, and their teachers, and school in general.

Charles Flowers (1966) investigated the effects of teacher expectancy on pupil achievement. Flowers used alleged ability grouping to establish teacher expectations. The study was conducted in two separate junior high schools located in two different cities. Both schools were in socio-economically depressed areas and the sample groups, two seventh-grade classes from each school (four groups in all), were educationally disadvantaged. The IQ levels and achievement levels in reading and arithmetic of all subjects were determined to be "average" and comparable to each other on the basis of a pretest. In each of the schools one of the two classes was allegedly of "high" ability; the other, serving as a control group, was cited as being of average ability. At the end of the school year, posttests were given to determine reading ability, arithmetic competence, and intellectual level (IQ). The results of the study indicated the existence of an educational self-fulfilling prophecy. In one school, the experimental (allegedly "high") group, when contrasted with the control group, showed a difference in achievement, although the difference was weak statistically; and there was no difference between the two groups in IQ. In the other school, the experimental group showed no difference

in achievement from the control group, but the difference in IQ was significant.

Rosenthal and Jacobson (1966) conducted a study with elementary school pupils. All the subjects were given a pretest, Flanagan's Tests of General Ability, which was disguised as a "test designed to predict academic 'blooming' or intellectual gain" (Rosenthal & Jacobson, 1966, p. 115). The school used a tracking system consisting of three levels--high, medium, and low. There were 18 classes in the six grades, three at each grade level. Within each of 18 classrooms, 20 per cent of the pupils were selected randomly and reported to the teacher as having the potential for unusual intellectual gains during the school year. A posttest was given eight months later, and significantly greater gains were made by the experimental pupils in the first and second grades. Only a selective summary of the literature has been presented, and the reader is referred to Rosenthal and Jacobson (1968) for an excellent review and discussion.

The nature of expectancy behavior. What are the behaviors of experimenters who are most influential in effecting prophesied responses in their subjects? What is the nature of the interaction between experimenters and subjects (between teachers and pupils) which communicates the expectations of one to the other? Upon examining the studies of communication of expectancy effects in experimental psychology, several variables have been investigated: sex, anxiety,

need for approval, acquaintanceship, experimenter status, and other related behavioral variables.

Rosenthal et al. (1964) found that male experimenters unintentionally bias the data collected from both male and female subjects. The female experimenters obtained significantly negative results with male subjects, indicating that perhaps male subjects over-react negatively to female experimenter influence. Fode (1965) found that there was a significant, but unpredictable, relationship between the anxiety level of either the experimenter or the subject and the magnitude of experimenter effects. Rosenthal (1966) and Crowne and Marlowe (1964) attempted to relate experimenter and subject need for approval to the degree of experimenter expectancy effects. They found that the nature of the relationship depends on the experimenter's level of anxiety. There was a significant positive relationship if experimenter anxiety was medium, and a negative relationship if experimenter anxiety was either high or low. Kanfer and Karas (1959) established a positive significant relationship between the acquaintanceship of the experimenter and the subject and the magnitude of expectancy effects.

In emphasizing the behavior of the experimenter in interaction with the subject, several interesting variables appear to be important. Rosenthal (1966) found that, even though the experimenter elicited certain expected effects unintentionally, he still manifested much the same behaviors

as those associated with more effective influencers who intentionally influenced the subjects. However, the process is still an extremely subtle one. Rosenthal (1966) also found that experimenters who exhibited greater expectancy effects tended to be perceived by their subjects as more professional and business-like; more relaxed (as opposed to nervous); speaking with a more expressive voice, and speaking more slowly; being more enthusiastic; and exhibiting fewer gross bodily movements, such as head and arm gestures and leg movements. If physical gestures were subtle, they did not detract.

Fode (1960), in an attempt to discover the nature of expectancy communication, placed a screen between the experimenter and subject. There was a distinct reduction in expectancy effects, indicating that visual cues from the experimenter are probably important. It must be noted, however, that the interposed screen did not entirely eliminate expectancy effects so that auditory cues are also important.

In a recent study by Adair and Epstein (1967) the experimenters' instructions to their subjects were tape-recorded. The self-fulfilling prophecy operated under both conditions, in the experimenters' presence and by tape-recording alone, indicating that expectancy effects can indeed be transmitted by the Prophet's voice.

There has been some research on the nature of teacher influence on pupil performance in typical classroom tasks,

but there have been few systematic observations. On the other hand, there is little research on the unintentional influence or communication which takes place as the result of teacher-pupil interaction in a typical classroom situation.

There are many variables involved in the communication of teacher expectancies, but one which is obvious is the possibility of the teacher's spending more time with the students who are to fulfill his positive expectations. However, Rosenthal and Jacobson (1968) reported that, according to teachers' judgments, no more time was spent with the experimental pupils than with the control pupils. The assessment was made subjectively and at the end of the school year. There was no systematic observation of time spent by the teachers with either group (Biddle & Adams, 1967).

Conn et al. (1967), in a study designed to test the ability of the subjects to judge the intent of a teacher by the tone of voice the teacher used, found that those pupils who could accurately interpret the meaning of the tone of the teacher's voice profited significantly greater from favorable teacher expectations than did those who were unable to interpret the voice tones. Thus, there is evidence that successful unintentional communication of teachers' expectations may depend upon various characteristics of the pupils themselves.

Interaction analysis. There are many different systems

for the analysis of interpersonal or group interactions, and the purpose for which the various techniques are designed range from counselor and client interaction to the interaction that occurs in a school classroom. Thus, there are only a few interaction analysis systems which are relevant to the purposes of this study, either in theory or technique. An observation system which was devised by Chapple (1949) emphasized the recording of the time element in interaction. According to Chapple, there should be a positive relationship between the amount of time two people spend in contact with each other and the quality of their attitudes toward one another and their emotional relationship.

Bales (1950) developed a system of interaction process analysis which was designed primarily to analyze communication in a wide variety of small problem solving groups. The objective of the system was described as follows: "It is assumed that the goal of the social scientist is to discover empirical generalizations about human behavior and to show that these observed uniformities are special cases or special combinations of more abstract and more general propositions" (p. 2). The general nature of the categories made the system unsuitable in its original form for use in the present study. However, Bale's theoretical framework and system of categories provided the structure for other systems and procedures closely related to this study.

Anderson (1945, 1946) was the first to apply a modi-

fied form of interaction analysis to classroom communication. Anderson and his associates were primarily interested in the dominative and integrative behavior of the teacher and the resulting effects on school children's behavior. Withall (1949) reduced the extensive number of categories which had been used in Anderson's system. As Medley and Mitzel (1962) point out, Withall (1949) studied classifier agreement, reliability, and validity in terms of relationship to Anderson's categories, pupil reactions, and expert ratings. In a study using the revised category system, Withall (1951) found that different teachers produce a different climate with the same group of pupils. Withall's sample, however, was extremely small. The technique as used by Withall was not a method for observing and recording behavior in the classroom; rather, it was a method for coding transcripts of sound recordings of classroom interaction.

Using Withall's categories, Mitzel and Rabinowitz (1953) categorized live teacher-pupil interaction, and were able to generalize their findings to the behavior of other teachers. Later Medley and Mitzel (1958) developed another technique for objectively observing and recording classroom behaviors. Their new instrument, the observation schedule and record (OScAR) was a modification of the techniques of Withall (1949) and Cornell (1952). Changes were introduced to increase observer reliability, to increase economy of observer time, and to refine the scoring process. Medley and

Mitzel (1958) concluded that relatively untrained observers using an instrument such as theirs can develop reliable information about differences in teacher-pupil interaction; and that the OScAR technique was sensitive to only three dimensions--verbal emphasis, emotional climate, and social structure.

Hughes (1963) developed a category system similar to Withall's except that Hughes and her colleagues observed non-verbal as well as verbal interaction. In one study, the data consisted of shorthand recordings by two observers of 129 half-hour sessions of the classroom behavior of four elementary children. At the end of the sessions the interaction was categorized and only those behaviors upon which the two observers agreed were recorded. As a result, Hughes and her associates obtained an optimum interaction pattern for teacher behavior. However, their conclusions were inconsistent with those of Medley and Mitzel (1958), who found that the variation in teacher behavior from observation to observation made such a model pattern ineffective.

The most sophisticated technique for recording observation of teacher-pupil interaction thus far has been developed by Flanders (1960). This system is particularly adapted to analyzing the influence pattern of the teacher--his control of the students' freedom to act. As Flanders (1960) states,

> Our interest is to distinguish those acts of the teacher that increase the students' freedom of

> action and to keep a record of both. The system of categories used forms a screen in front of the observer's eyes so that those acts which result in compliance are sharply separated from those that invite more creative and voluntary participation, while certain aspects of subject matter are ignored. . . . Interaction analysis is (therefore) a process of abstracting the intent of an act from the act itself (pp. 1-2).

Interaction analysis, therefore, attempts to record only verbal interaction, the assumption being that the verbal behavior of an individual is an adequate sample of his total behavior and that verbal behavior seems to correlate highly with non-verbal behavior.

Statement of the Problem

There has been little systematic attempt to observe the way in which teachers communicate their expectations about pupil achievement to their pupils. The present study was designed to explore the means by which this is accomplished and to determine the extent of the effects of teacher expectations on pupil attitudes toward school-related concepts and pupil achievement in the language arts and mathematics.

The problem stated in terms of formal hypotheses is as follows:

1. Pupils who are expected by their teachers to achieve well will perform better in the subjects tested than those pupils of like ability of whom nothing special is expected.
2. Pupils who are expected by their teachers to

achieve well will develop attitudes which are more favorable toward school-related concepts than those of like ability of whom nothing special is expected.

3. Teachers will communicate in a more positive-accepting-supportive manner to the experimental pupils than to the control pupils. (Positive-accepting-supportive behavior is defined as that behavior which is recorded in categories 1, 2, or 3 of the Category System of Observation.)
4. Teachers will communicate in a more negative-rejecting-defensive manner to the control pupils than to the experimental pupils. (Negative-rejecting-defensive behavior is defined as that behavior which is recorded in categories 5, 6, or 7 of the Category System of Observation.)
5. The magnitude of teacher expectations as depicted in pupil gains in achievement is greater in the language arts than in mathematics.
6. Teachers spend more time communicating with pupils whom they consider to be "bright" than with pupils of like ability whom they consider to be "average."
7. Teachers communicate more often with pupils whom they consider to be "bright" than with pupils of like ability whom they consider to be "average."
8. The degree to which a teacher's expectations influence pupil achievement and attitudes is a func-

tion of the amount of time the teacher spends communicating with the pupil.

9. The degree to which a teacher's expectations influence pupil achievement and attitudes is a function of the number of occurrences of teacher communication to the pupil.

10. Positive-accepting-supportive pupil communication to teacher increases the magnitude of positive-accepting-supportive communication of teacher to pupil.

CHAPTER II

METHOD

The first aspect of the present study was devoted to identifying how teacher expectations were communicated and involved a series of four one-hour sessions of observation of the interaction between 23 teachers and 75 seventh-grade experimental pupils and 75 seventh-grade control pupils. The second part of the experiment was concerned with changes in pupil achievement and attitudes as the result of teacher expectations about pupil performance. The experimental design used was Random Replication (A X R) of Treatments X Subjects and Treatments X Levels (Lindguist, 1953).

Pre-experimental Procedures

Selection of subjects. The sample was drawn from six middle-class junior high schools in three different Oklahoma cities, and consisted of an experimental group of 75 seventh-grade pupils of "average" ability, a control group of 75 seventh-grade pupils of "average" ability, and 23 English and mathematics teachers. "Average" ability was defined as that ability indicated by a score of 90 to 110 on the Otis-Lennon Mental Abilities Test: Intermediate Level.

During the summer preceding the fall school term, the cumulative records of all seventh-grade pupils in the six junior high schools were screened, and on the basis of available scores on various intelligence tests which had been given in grades four to six, pupils were selected as potential subjects.

In scheduling these pupils, two criteria were followed. First, the pupils were assigned to English and mathematics sections designated as "average." All participating schools followed practices of grouping pupils according to "ability level." Second, an experimental or control pupil in any one class, say English, could not appear in a mathematics class which was being used in the experiment. Thus, no subject was exposed to the experimental treatment by more than one teacher. As a consequence of the selection procedure, only about 300 pupils of a total seventh-grade population of some 1500 were ultimately potential subjects. The final determination of subjects who met the criterion of "average ability" as operationally defined for this study was made at the time the pretests were given.

Selection of instruments. Four instruments were utilized to collect the data: (1) The Category System of Observation (CSO) for recording teacher-pupil communication, and which was essentially a modification of Bales' Interaction Process Analysis and Chapple's technique for measuring interaction on a temporal basis; (2) the Stanford Achievement

Tests: Partial Advanced Battery (Language Arts and Mathematics); (3) the Otis-Lennon Mental Abilities Test: Intermediate Level; and (4) semantic differential scales.

Systems of interaction analysis. Although there were many systems of analyzing interpersonal or group interactions, the purposes for which the various techniques were designed ranged from recording counselor and client interaction to analyzing the interaction that occurs among all members of a school classroom. Those systems which were developed specifically to record classroom interaction (Flanders, 1964; Hughes, 1963; Medley & Mitzel, 1958; Withall, 1951) emphasized the interaction of the total group and focused rather directly on the cognitive aspects of the teaching-learning process. As such, they would have generated data superfluous to this study. The study of teachers' communication of their expectations required an instrument in which the focus was upon those variables which were predicted to be particularly relevant, and none of the existing techniques were especially suitable.

Bales (1950) developed a category system of interaction analysis which was designed primarily to record communication in a wide variety of small problem solving groups. Although the general focus of Bales' system was not appropriate for the needs of this study, some of the category definitions and the general structure of the system were suitable and provide the basis for the design of the CSO. In addition, an

observation technique devised by Chapple (1949) emphasized the recording of the time element in communication. According to Chapple, there should be a positive relationship between the amount of time two people spend in contact with each other and the quality of their interpersonal communication. Although Chapple's recording procedure was not employed, his theory provided the rationale for using time as well as categories to measuring interaction quality.

Rosenthal and Jacobson (1968) have identified three major dimensions of possible importance to the self-fulfilling prophecy phenomenon in the classroom: first, the positive-negative nature of teacher-pupil interaction; second, the amount of time or frequency of teacher communication to a pupil; and, third, whether teacher communication was verbal or non-verbal. The CSO combined these three rather broad dimensions into a system of seven categories specifying the quality of behavior observed and four categories which specify whether the communication is from teacher to pupil or pupil to teacher, and whether the pupils communicating were of the experimental or control group. The amount of time spent in each instance of communication is used as the unit of measure.

Three categories were used to record the positive-accepting-supportive communication of the teacher to his pupils and the pupils to their teacher. Category 1 provided for the measurement of direct, verbal, positive references to the pu-

pils' ability level (e. g., "You're too bright to be wasting your time"). Category 2 was used to record measurements of communication of a positive, verbal nature, but which are not direct references to the pupils' ability level (e. g., praise, extolling, concurrence, kidding). Category 3 provided for communication of a positive but non-verbal nature (e. g., smiling, waving, admiring, nodding approval).

That behavior which was purely task oriented and therefore neutral (e. g., giving information, clarifying, or repeating with no discernible tendency toward positive or negative behavior) was recorded in Category 4.

In addition, there were three categories in which negative-rejecting-defensive communication of the teacher to his pupils and of the pupils to their teacher was recorded. Category 5 was used to record all non-verbal, negative communication (e. g., cool, detached, aloof, inattentive). In Category 6 the observer recorded all verbal communication of a negative nature which did not make specific reference to the pupil's ability to perform in school (e. g., sarcasm, nagging, scoffing, extreme arbitrariness). Category 7 was used to record direct, verbal references to the pupils' lack of academic ability (e. g., "I'm no genius," or "You may not be the smartest student in the world, but . . ."). The categories for the system of observation used in the present study are shown in Figure 1 and are defined in Ap-

pendix A.

In order to score the CSO, the observer was seated so that he could perceive all teacher-pupil interaction, both verbal and non-verbal. Since communication can be established between two individuals even though one of them is engaged in interacting with a third person, the observer was constantly attentive to facial expressions and bodily movements as well as verbal intercourse. Verbal and non-verbal expressions which could not be perceived by the person toward whom the expression is being directed were not recorded. The communication had to have a sender and a receiver. In addition, the observer had to be able to distinguish among teacher-experimental group interaction, teacher-control group interaction, and teacher interaction with other members of the class.

As was previously mentioned, all interaction between the teacher and either the experimental pupils or the control pupils was recorded in a category indicating the type of communication and whether the interaction is teacher initiated or pupil initiated. Each entry was recorded as a measure of time (in seconds) and, thus, provides both a measurement of the frequency of communication and the amount of time spent communicating.

Form J of the Otis-Lennon Mental Abilities Test: Intermediate Level (Appendix C) was used to determine the ability levels of the pupils. This instrument was selected for two

1	Direct Reference (Positive) --Verbal--
2	Accepting/Supportive --Verbal--
3	Accepting/Supportive --Non-Verbal--
4	Task Oriented (Neutral)
5	Rejecting/Defensive --Non-Verbal--
6	Rejecting/Defensive --Verbal--
7	Direct Reference (Negative) --Verbal--

Fig. 1. Categories of the Category System of Observation

reasons: (1) It offered an adequate estimate of the general range of intellectual functioning of a pupil at the seventh-grade level; and (2) it made use of the same answer sheet as used in taking the Stanford Achievement Battery. The Stanford Achievement Tests: Partial Advanced Battery (Form X) was selected to measure pupil achievement in the two subject areas being tested, language arts and mathematics. The instrument has been well standardized and requires minimum administration time (see Appendix D).

Two sets of semantic differential (Osgood, 1957) concepts and scales were used--one as a measure of change in attitudes of pupils, the other as a posttest measure of teacher attitudes. The concepts on the semantic differential for the pupil were "Teacher," "Me," "School," "Fellow Student," "Math," and "English." The teacher was measured on the concepts of "Teaching," "Pupils," "Tests," "Achievement," "Intelligence," and "Mediocrity" (see Appendices E and F). In addition, a brief questionnaire (Appendix G) was administered to all teachers to determine if at any time they gained knowledge of or insight into the real purpose of the experiment.

All instruments except the CSO had been thoroughly field tested or standardized and were considered to possess adequate validity and reliability. The CSO appeared to be near enough to the structure of Bales' system and to the principle of operation of Chapple's system that the reliability

and validity of those techniques was believed to lend credence to the CSO, at least for the purposes of the research for which it was constructed.

Although observer reliability can be established in a number of ways, the emphasis was on observer agreement in the present study. The investigator periodically had all observers record the same teacher-pupil interaction. An estimate of reliability was then determined by Scott's (1955) coefficient, "Pi." Scott's method is "unaffected by low frequencies, can be adapted to percentage figures, can be estimated more rapidly in the field, and is more sensitive at higher levels of reliability" (Flanders, 1960, p. 13). Reliability coefficients of teacher to pupil communication using amount of time as the measure ranged from .85 to .95. Reliability estimates of pupil to teacher communication using the same unit of measure ranged from -.12 to +.69, as training time on the CSO increased. Therefore, as with most category systems, extensive training of observers was required if reliable data was to be collected.

Experimental Procedures

Establishing teacher expectations. Expectations regarding pupil abilities were established by a two-phase communication. During the teacher preparation week before school started, meetings were held with the English and mathematics teachers in each participating school, and the teachers were given the following verbal communique:

Good morning/afternoon. I am Mr. Kester, Assistant Professor of Education at Oklahoma Baptist University. I need to use some of the pupils scheduled in one of your seventh-grade classes for an experiment related to my dissertation. It is anticipated that it will require a minimum of interruption in that class.

In general, the experiment is related to the manner in which "intellectually superior" pupils relate to each other and to peers of "average" intellectual ability within an "average" setting. The theoretical construct involved is the social intelligence identified by J. P. Guilford (said parenthetically). We have placed a few "above-average" pupils in your "average" sections of English or math. They were tentatively selected on the basis of IQ scores available in their cumulative records. During a two-day testing session next week, the ability level of these pupils will be confirmed and you will be notified as to their identity.

The students will be taken from your ___ hour class (designate to each teacher which of his classes will be used in the experiment) during the first two days of school for some tests. They will be tested again in nine weeks. During the testing sessions, other pupils in your classes will also be tested. This measure is intended to provide camouflage to prevent the experimental subjects from realizing that they are of special interest to us (to alleviate the "guinea pig" effect, as it is called). At no time during the entire experiment are the "bright" pupils to be aware that they are experimental subjects. In addition, during the second, fourth, sixth, and eighth weeks, an observer would like to come into your class and observe the social interaction of our experimental subjects with their "average" peers.

Because experimental error in typical educational field studies is introduced from many sources, I would like to ask that, if you agree to assist me, from this moment until the end of the experiment you talk to no one about the experiment--neither to pupils, fellow teachers, counselors, administrators, nor my observers. Even though the counselors and administrators know the full details of the experiment, someone would invariably overhear your discussion and jump to some incorrect, and conceivably injurious, conclusions.

Are we agreed?

> You will receive an additional communique after the initial testing sessions, and the observer will contact you before the observation session to locate his position in the classroom and ask you for a class seating chart. Other than that (and your keeping "mum"), you'll not have any part in the experiment.

Are there any questions?

After pretests had been administered and the subjects assigned to experimental or control groups, teachers were given a written communique (Appendix G) which listed the names of the "bright" pupils and reiterated that the nature of the experiment was to observe the manner in which such "bright" pupils behaved in a classroom situation among "average" peers. The teacher did not receive any information regarding the control pupils; in fact, the teachers were supposedly unaware that there were control subjects in the study. Again the teachers were asked to refrain from discussing the experiment with anyone in order to "control the 'guinea pig' effect in pupil behavior." (This instruction was actually an attempt to control the exchange of information among teachers which might lead them to suspect the true nature of the experiment.) In addition, the cumulative records of the experimental pupils were pulled from the school files and were not accessible to teachers during the period in which the experiment was being conducted.

Administration of the pretests. During the first week of school, all the potential subjects were given the Stanford Achievement Tests (Language Arts and Mathematics), the

Otis-Lennon Mental Abilities Test, and the pupil's semantic differential scales. Upon receiving the results of the Otis-Lennon, subjects within each class to be used in the experiment, those who scored "average" or 90-110, were randomly assigned to either the experimental group or the control group in one of the English or mathematics classes. The experimental group in each class was the same size as the control group. In some of the classes, the experimental group and control group were as small as two subjects; in others, they were as large as six. As a result of the final selection of subjects during the pretest, 75 seventh-grade experimental pupils and 75 seventh-grade control pupils were used in the study.

Observations of teacher-pupil communication. During the second, fourth, sixth, and eighth weeks of school, an observer visited each class containing the matched groups of pupils and, using the CSO, recorded teacher-pupil and pupil-teacher interaction for both the experimental and control groups. Each observation session was fifty minutes in length. The total observation time for all experimental classes was 80-100 hours in four successive measurements over a period of eight weeks.

Administration of the posttests. During the ninth week of school the same instruments used in the pretests were readministered to both the experimental and control groups. In addition, teachers were given a semantic differential

measurement (Appendix F) and the questionnaire related to their knowledge of the nature of the experiment (Appendix G).

CHAPTER III

RESULTS

Data recorded for pupil subjects (S_p) included pretest scores for mathematics achievement, English achievement, IQ, and attitudes; posttest scores for the same areas; and difference scores for each area, to include the direction of any change. The data of teacher subjects (S_t) included teacher to experimental pupil communication (T - P_E); teacher to pupil control group communication (T - P_C); pupil experimental group to teacher communication (P_E - T); and pupil control group to teacher communication (P_C - T). Teacher subjects' data also included a posttest only measure of attitude.

The communication data were originally measured in seconds, but were converted to tenths of minutes when recorded. Each instance of communication was also recorded as frequency data. Thus, teacher-pupil interaction was recorded as a time and occurrence measurement according to category (see Figure 1).

The semantic differentials consisted of six concepts, each measured by seven different seven-point scales. The

adjective pairs were heavily weighted with the evaluative factor. Each scale was scored from one to seven, a score of one being the lowest (nearest the undesirable or negative adjective).

Where analysis of variance was the statistic used to analyze the data, an integrated technique which was developed by A. E. Dahlke (1966) was employed. In addition, references to "occurrences" of communication appear to imply frequency data. However, occurrence data were transformed into means within each category, and summed across teacher subjects. They were, therefore, continuous data, and analysis of variance was an appropriate statistic.

Analysis of Teacher-Pupil Communication

Measurement of teacher-pupil communication was made in several dimensions: the quality of the interaction as indicated by the accepting or rejecting, friendly or unfriendly nature of the communication; the verbal or nonverbal nature of the communication; the amount of time spent communicating; and the occurrence of communication.

The quality of interaction. The primary concern of the present study was to investigate the quality of the communication between teacher and pupil and between pupil and teacher. There were seven categories of communicative behavior in which the amount of time of each instance of communication was recorded. As described in the previous chapter on methods and in Appendix A, there were three cate-

gories in which positive-accepting-supportive communication of the teacher to his pupils and the pupils to their teacher was recorded. Category 1 provided for measurements of direct, verbal, positive references to the pupils' ability level. Category 2 was used to record measurements of communication of a positive, verbal nature, but which were not direct references to the pupils' ability level. Category 3 provided for communication of a positive but non-verbal nature.

There were also three categories in which negative-rejecting-defensive communication of the teacher to his pupils and of the pupils to their teacher was recorded. Category 7 was used to record direct, verbal references to the pupils' lack of academic ability. In category 6 were recorded all other verbal communication of a negative nature. Category 5 was used to record all non-verbal, negative communication. Category 4 was used to classify all communication which was completely task-oriented and was indiscernible as to its positive or negative nature.

Hypothesis 1 stated that teachers would communicate in a more positive, accepting, supportive, friendly manner to pupils whom they thought to be intellectually superior than they would to pupils of like ability, but whom they considered to be average. A 2 X 4 (groups X replications) analysis of variance with repeated measures in the replications dimension comparing the positive-accepting-supportive

communication of teachers to the experimental and control pupils was significant at the $p < .001$ level, using the amount of time as the measure (see Table 1). When using occurrence measures, the difference was significant at the $p < .05$ level, between the experimental and control groups (see Table 2). In addition, the amount of time spent by the teachers communicating to the pupils in a positive-accepting-supportive manner increased significantly with each measure ($p < .05$; Table 1).

The present experimenter also predicted that teachers would communicate in a more negative, rejecting, and defensive manner toward their pupils of average ability than toward the allegedly bright pupils. However, the difference in the teachers' negative communication toward each group was not significant.

Table 1

Analysis of Variance Performed on the Time Teachers Spent Communicating in a Positive-Accepting-Supportive Manner

Source	df	MS	F	p
Groups	1	676.26	17.16	.001
Measures	3	53.12	3.31	.05
Interaction	3	9.30	.58	NS
Error [S(R-1)]	44	39.41		
Error [S(R-1)(M-1)]	132	16.03		

Table 2

Analysis of Variance Performed on the Occurrences of Teachers' Communicating in a Positive-Accepting-Supportive Manner

Source	df	MS	F	p
Groups	1	2724.26	6.54	.05
Error	44	416.81		

Interaction time and occurrences. In order to assess the amount of time spent by the teacher communicating to the experimental pupils as opposed to the control pupils, a 2 X 4 (groups X replications) analysis of variance with repeated measures over the replications was used. The teachers in the study spent significantly more time ($p < .005$) communicating with pupils whom they considered to be "bright" than with the control pupils of like ability, whom they considered to be "average." (See Table 3.)

Comparing the occurrences of teacher to experimental pupils as opposed to control pupils over the four measurements was accomplished by a 2 X 4 (groups X replications) analysis of variance with repeated measures over the replications dimension. However, when using occurrence data, no significant difference between the groups was found.

The relationship of P - T to T - P interaction. In order to test the assertion that positive-accepting-supportive pupil communication to teacher would increase the magnitude of positive-accepting-supportive teacher communica-

tion to pupil, the time teachers spent in positive communication was divided into three levels of pupil communication to teacher--high, medium, and low. Thus, teacher communication to experimental and control pupils was compared to the level of pupil to teacher communication. A 2 X 3 (groups X levels) analysis of variance yielded no significance for the main effect of groups. However, a linear trend for the experimental group's data indicated significance at the $p < .001$ level (see Table 4).

Analysis of Expectancy Effects on Pupil Achievement

Achievement resulting from teacher expectations. The importance of the self-fulfilling prophecy in the study was the effect that teacher expectations had upon pupil achievement in the two subject areas measured--mathematics and language arts. It was hypothesized that pupils who were expected by their teachers to achieve well would perform better in the subjects tested than those pupils of like ability of whom nothing was expected. A 2 X 4 (groups X replications) analysis of variance with repeated measures on the replications dimension yielded no significant difference in pupil gain in language arts as opposed to mathematics ($p < .025$; see Table 5).

The investigator also suspected that the more "subjective" interaction of the language arts classroom would provide a more functional climate in which teacher expecta-

Table 3

Analysis of Variance Performed on the Amount of Time Teachers Spent Communicating with Experimental Pupils vs. Control Pupils

Source	df	MS	F	p
Groups	1	588.27	11.17	.005
Measures	3	83.01	4.32	.05
Linear Trend	1	247.31	4.78	.05
Interaction	3	8.45	.44	NS
Error [G(R-1)]	44	51.76		
Error [G(R-1)(M-1)]	132	19.20		

Table 4

Linear Trend Analysis Performed on the Relationship Between Positive P-T Communication and Positive T-P Communication

Source	df	MS	F	p
Linear Trend	1	4199.10*	16.37	.001
Error	20	256.46		

*Corrected for unequal N

tions could be communicated than the more "objective" interaction of the mathematics class. The hypothesis was, then, that the effect of teacher expectations as depicted by pupil gains in achievement would be greater in the language arts than in mathematics. As reported above, there was significant difference between pupil performance in the two subjects. However, the difference noted may not necessarily be due to the experimental effect. Contrasts comparing the experimental group versus the control group for language arts and for mathematics indicated no significance.

Table 5

Analysis of Variance Performed on the Gain in Pupil Achievement in Language Arts as Opposed to Mathematics

Source	df	MS	F	p
LA X M	1	1192.01	5.77	.025
Error	148	206.25		

Pupil achievement as a function of time and occurrences. In analyzing the effect of the amount of time or the occurrences of teacher communication on pupil achievement, both experimental and control pupils' achievement scores in mathematics and language arts were grouped either above or below the median on T - P communication. A 2 X 2 (groups X above-below medians) analysis of variance for both mathe-

matics and language arts indicated no significant relationship between the amount of time or the occurrences of teacher communication and the gain in pupil achievement in those subject areas.

Analysis of Pupil Attitudes

Teacher expectations and pupil attitudes. It was proposed that if a pupil was expected to achieve well, he would also develop attitudes which are more favorable toward school, his teacher, himself, the subjects which the experimental teachers taught (mathematics and English) and his fellow students than those pupils of like ability, but of whom nothing special was expected. The attitude change scores for each concept or object were obtained from the semantic differential scales. A 2 X 6 (groups X attitudes) analysis of variance comparing the attitudes of the experimental group to those of the control group was not significant.

There was, however, a significant difference ($p < .01$) among the attitudes of the pupils (see Table 6). A contrast between language arts and mathematics was significant at the $p < .001$ level, thus accounting for approximately 70 per cent of the variance obtained when measuring the difference among all attitudes (see Table 6).

Pupil attitudes as a function of time and occurrences. In analyzing the degree to which a teacher's expectations influence pupil attitudes, the attitude scores for the experimental pupils and for the control pupils were catego-

rized as above the median or below the median on teacher to pupil communication time and occurrence. A 2 X 2 (groups X above-below medians) analysis of variance matrix was set up for the time teachers spent communicating and the frequency of teacher communication. Neither analysis yielded significance.

Table 6

Analysis of Variance Performed on Pupil Attitudes Toward Teacher, Self, School, Mathematics (M), English (LA), and Fellow Students

Source	df	MS	F	p
Attitude Measures	5	384.70	9.17	.01
Contrast (M X LA)	1	1395.36	33.27	.001
Error	740	41.93		

Analysis of Other Data

Teacher expectations and IQ change. Although there was no hypothesis related to the effect of teacher expectations on pupil IQ, the data were collected and analyzed for purposes of future research. It was not surprising to find that a 1 X 2 (groups) analysis of variance comparing the IQ gain of the experimental pupils with the IQ gain of the control pupils showed no significant difference.

Teacher attitudes. As related in the methodology chapter, a posttest measure of the attitudes of S_t's toward the

concepts of "teaching," "pupils," "tests," "achievement," "intelligence," and "mediocrity" was obtained. The data were analyzed using the D statistic. The D matrix for the data is given in Table 7. Three interesting clusters of concepts are evident: the first consists of "teaching," "pupils," "achievement," and "intelligence"; the concept "tests" stands alone; likewise, the concept "mediocrity" stands alone, and is considerably distant from all other concepts rated (D's = 3.40-- 6.02). In addition, the concepts of "achievement" and "intelligence" were extremely close together in semantic space (D = .69).

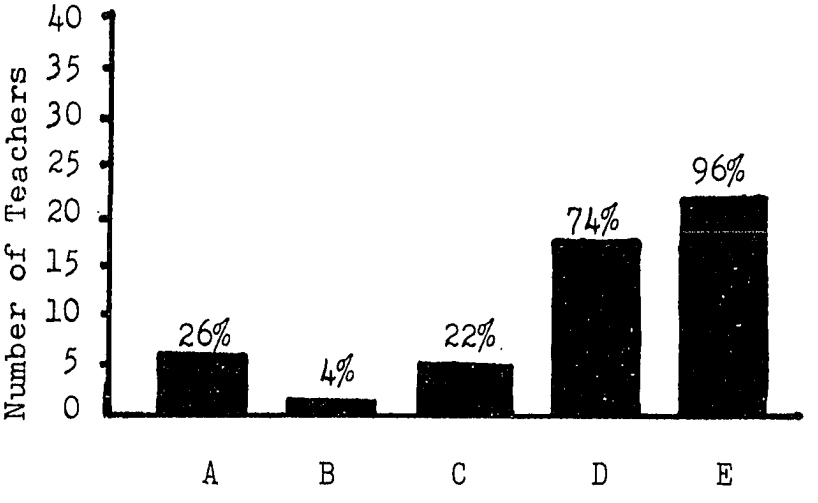
Table 7

D Matrix for Teacher Attitudes

Concepts	Tchg	Pupl	Tsts	Achv	Intell	Medioc
Tchg		.92	3.01	1.15	1.74	6.02
Pupl	.92		2.89	1.05	1.44	5.62
Tsts	3.01	2.89		2.44	2.15	3.40
Achv	1.15	1.05	2.44		.69	5.43
Intell	1.74	1.44	2.15	.69		4.97
Medioc	6.02	5.62	3.40	5.43	4.97	

Teacher awareness. In response to the end-of-study questionnaires regarding the teachers' knowledge of the true nature of the experiment, only six teachers reported that

they suspected that the real purpose of the experiment was different from that which they were originally led to believe. Of those six, only one correctly identified the true nature of the study. Thus, for all practical purposes, 22 of the S_t's had no knowledge of the experiment which might alter the effect of the experimental condition (see Figure 2).



Key:

A - Teachers who suspected that the real purpose of the experiment was different than they were told (6).

B - Teachers from category A who were correct in identifying the real purpose of the experiment (1).

C - Teachers from category A who incorrectly identified the real purpose of the experiment (5).

D - Teachers who did not suspect that the real purpose of the experiment was different than they were told (17).

E - Teachers who did not know the real purpose of the experiment (22).

Fig. 2. Histogram of the Teachers' Knowledge of the Experiment

CHAPTER IV

DISCUSSION

The basic assumption underlying the present study was that a prophecy can be instrumental in its own fulfillment (Merton, 1948). Specifically, the self-fulfilling prophecy concept was applied to an educational setting at the secondary school level. The major purpose of the study was to investigate how teachers communicate their expectations about pupil performance to the pupils, and to determine the effects of such teacher expectancies on pupil achievement in mathematics and language arts and on pupil attitudes toward certain school-related concepts.

Teacher Communication of Expectations

Initiating the experimental condition. The logic used for establishing teacher expectations was essentially the same as was used by Pitt (1956), by Flowers (1966), and by Rosenthal and Jacobson (1968)--a fictitious ability level for selected pupils was reported to the teachers involved in the experiment. The experimental teachers were told that the selected pupils were "bright" (IQ greater than 120). None of the teachers had had previous contact with the pupils.

In addition, the teachers were told that these "bright" pupils had been placed in one of their "average" classes in order to observe the bright pupils' interaction with their average peers. They were instructed not to treat the pupils differentially.

Of the 23 teacher subjects, 17 evidently suspected no deviation from the purpose of the experiment as it was originally explained to them. Of the six who did suspect that the purpose might be different, only one suspected that her behavior was being observed to see if she behaved in a special way toward the bright pupils. None were aware of the observation of their interaction with a control group of pupils.

The dimensions of communication quality. How does a teacher communicate his prophecy to his pupils? What is the nature of the teacher-pupil interaction which may lead to the self-fulfillment of the prophecy? Rosenthal (1966) noted some individual differences among experimenters in the degree to which they obtained results consistent with their expectations. Those experimenters who were most effective in communicating their expectations were more friendly, likeable, and encouraging, but not overly personal. The process of communicating one's prophecies also appears to be unintentional and very subtle. Both visual and auditory cues were shown to be operative in studies done by Fode (1960). In the present study, the teachers' communica-

tion to both the experimental and control pupils, and the communication of both groups of pupils to the teachers was recorded in the categories of the CSO as both a time measure and an occurrence measure. The rationale for the use of amount of time was originally proposed by Chapple (1949). It was Chapple's contention that there is a positive relationship between the amount of time two people spend in contact and the quality of the relationship between them. Although the amount of time spent communicating is significantly, positively correlated with the occurrences of communication, amount of time appeared to provide a better index of the amount of contact. An additional consideration was Flanders' (1965) assumption that the verbal behavior was an adequate sample of a person's total behavior and was highly correlated with nonverbal behavior. The CSO, therefore, provided not only a means by which to measure the quality of behavior, but also as the quality is verbally and nonverbally communicated and as it is indicated by the amount of time spent in communication.

The quality of teacher-pupil communication. Considering both the verbal and non-verbal behavior, teachers tend to communicate in a significantly more encouraging, supportive, friendly manner toward pupils whom they consider to be intellectually superior than they do toward average pupils. Since both P_E's and P_C's were taught by the same teachers, the differential treatment was not due to teacher differ-

ences, as might well have been the case in Flowers' (1966) study. There are, of course, several possible explanations for the variability which teachers displayed in the quality of their communication.

One can deduce that the experimental condition was operative, and that the teachers tended to give preferential treatment to those pupils who were supposedly bright as opposed to average or below average. It is also possible that the scores yielded by the CSO were but an artifact of the instrument itself. The construct validity of the instrument warrants additional study. Too, perhaps the self-fulfilling prophecy was operational in observer behavior. The observers knew which pupils composed the experimental group and which composed the control group. Thus, even though observers reported a conscious attempt not to succumb to experimenter effects, observer bias is an obvious possibility.

Another distinctly possible explanation is that after even a few days of classroom contact, the teachers were able to tell that the "bright" pupils were not performing better than some others in the class who were only average. Such was reported by four teachers and eventually led them to suspect that the purpose of the experiment might be different from that which was originally stated. Under these conditions, one might expect the teacher to increase his efforts to get the allegedly bright child to perform up to his

ability level. A linear trend of teacher communication over observations was significant, and would seem to support such a hypothesis.

An additional interpretation can be made--an interpretation which does not preclude the operation of the experimental condition. One might suspect that the typical middle-class, white teacher might identify, or attempt to identify, more readily with brighter pupils. Brighter pupils generally have a higher level of performance which in turn can serve as a reinforcement for the teacher's effort.

One of the early criticisms of interpreting Rosenthal's (1966) results as evidence of the existence of the self-fulfilling prophecy phenomenon was that such results could possibly be accounted for in terms of operant conditioning. It was thought that perhaps the experimenters were shaping the behavior of their subjects by unintentionally and subtly rewarding them when they responded as the experimenter expected them to respond. Rosenthal and Jacobson (1968) point out that if that were the case, one would expect a typical learning curve for the subjects. However, no learning curve was found. In fact, the subjects' first responses were affected as much as their last responses (Rosenthal, 1966; Rosenthal & Jacobson, 1968).

There did, however, appear to be a learning curve for experimenter behavior (Rosenthal, 1966). Rosenthal and Ja-

cobson (1968) hypothesized that a subject's responding in the direction predicted by the experimenter is reinforcing to the experimenter, and that subjects, therefore, may unintentionally shape experimenter behavior. The results of the present study do appear to confirm this latter contention. As the magnitude of positive, supportive, accepting communication of the pupils toward the teacher increases, there is a significant tendency for the magnitude of the positive, supportive, accepting communication of the teacher to the pupils to increase. There is an unintentional shaping of behavior, but it is the pupils who shape the teacher's behavior in the process of teacher expectancy communication.

The amount of teacher-pupil interaction time. Ten months after the posttest, Rosenthal and Jacobson (1968) asked the teachers in their experiment to estimate the amount of time they had spent with both the experimental pupils and the control pupils. There was no significant difference in the time spent with one group or the other. Rosenthal and Jacobson (1968) concluded that possibly direct observation of teacher-pupil interactions might have given different results, but that it seemed plausible to think that it was not a difference in the amount of time spent with the children of the two groups which led to the differences in their rates of intellectual development, but rather the quality of the interaction.

In view of the lapse of time between the end of the

Rosenthal and Jacobson (1968) study and their survey of teacher opinion as to the amount of time spent, and in the light of Chapple's (1949) assumption that the amount of time spent interacting and the quality of interaction are one and the same, direct observation of the time spent in teacher to pupil communication was investigated. The teachers of the present study spent more time communicating with their allegedly brighter pupils, and there appeared to be a high positive correlation between accepting, encouraging, friendly behavior and the amount of time spent by the teacher in communication--a finding which substantiates Chapple's contention. Even so, when the difference in the amount of time the teachers spent communicating with those pupils whom they thought were bright was compared to the level of pupil achievement, there was no meaningful difference between the experimental and control groups. Thus, even though a teacher may spend more time with his allegedly "brighter" pupils, that evidently affected neither pupil achievement nor intellectual ability, at least not within a nine-week period of time and not after the pupils were already in the seventh grade.

The Effects of Teacher Expectations on Pupil IQ and Achievement

The most dramatic results in experimentation on interpersonal self-fulfilling prophecies have been obtained by

Rosenthal and Jacobson (1968). The first and second grade experimental pupils in their studies showed significant gains in IQ over the control pupils. Their fifth and sixth grade experimental pupils showed gains over the control pupils on language and arithmetic which approached significance. Pitt (1956), on the other hand, found no significant difference between his fifth grade experimental and control pupils in objective test scores or school grades. Charles Flowers (1966), using seventh grade pupils in two different schools, did find that his allegedly superior group in one school obtained higher, but statistically insignificant, scores in reading and arithmetic than its control class, and that the experimental group in his other school made significant gains in IQ over its control class. In summary, then, experimentation to date indicates that the effects of the teacher's expectations about pupil achievement based on a knowledge of the pupil's intelligence level is an elusive phenomenon. There is also considerable evidence that it is unlikely that teacher expectations about a pupil's intelligence level is significant, except in the early grades of school.

In the present study, no significant gains in language arts or mathematics achievement by the P_E's over the P_C's were found. When pupil gains in language arts as compared to mathematics were analyzed, there was, however, a significant difference in favor of the language arts. It was

predicted that the highly verbal, relatively subjective climate of the language arts classroom would provide a more functional medium in which the teacher could communicate his expectations to his pupils. However, as previously reported, an analysis of the experimental group versus the control group for each subject area showed no significant differences. Therefore, it is doubtful that there is a meaningful difference in the interaction climate of a classroom that can be attributed to the subject area itself. It is also noteworthy that there was no significant differentiation between the IQ gains of the experimental group and those of the control group in this study.

There is one point of possible importance in interpreting the lack of achievement and IQ gains in this experiment. Examiners reported that all pupils were extremely concerned about their missing two days of classes required for each testing session, especially in order to take tests which the other pupils were not taking and which would not count as part of their school record. In addition, although standardized instruments were used and examiners were given what was considered to be sufficient training in administering them some variation in testing procedures is suspected. Thus, it is probable that an intensified examiner training program would have increased the reliability of the scores obtained.

The Effects of Teacher Expectations on Pupil Attitudes

Previous experimentation indicates that teacher expectations about pupils' ability levels may have a meaningful effect on the pupils' attitudes and interests. Some of Rosenthal and Jacobson's (1968) special pupils were assessed by their teachers as being more intellectually curious; more interesting, appealing, and happy; and possessing a greater probability of future success. Pitt (1956) also found that the teachers' beliefs about pupils' IQ levels affected the pupils' views of themselves, their teachers, and school.

In an effort to corroborate Pitt's and Rosenthal and Jacobson's findings, which were based on subjective analyses of teacher and pupil ratings, semantic differential scales were used to measure several school-related concepts: school, teacher, self, mathematics, English, and fellow students. There appeared to be no significant difference in the experimental pupils' changes in attitudes as compared to those of the control group. However, it is very likely that the lack of simplified directions for marking the rather complex scales could have seriously affected the reliability of the attitude measures.

Some interesting results were, nevertheless, obtained. There was significant variation in the attitudes of all pupils, particularly between their attitudes toward mathematics and English. The positive change pupil attitudes toward English was highly significant when compared to the

positive change in their attitudes toward mathematics.

Teacher Attitudes

In a questionnaire approach, Flowers (1966) compared the attitudes of the teachers of his two experimental classes to those of his control classes. He found that the teachers of the experimental groups preferred to teach their "bright" classes; that their allegedly superior classes offered fewer disciplinary problems, and that they had more positive attitudes toward the teaching process and pupil accomplishments in the experimental classes. In an attempt to systematically analyze attitudes toward certain concepts related to the present study, a semantic differential schema was administered to the teachers after the posttest data had been gathered on the pupils. A pretest-posttest design would have been more useful, but that would have seriously contaminated the experimental condition.

Teacher attitudes toward teaching, pupils, tests, achievement, intelligence, and mediocrity revealed some interesting relationships. By examining the means of the various concepts and their distance in semantic space, clusters appeared. Teaching and pupils were very close together in meaning, as were intelligence and achievement. In fact, an analysis using the D statistic showed intelligence and achievement to be closer together than any of the other concepts. It appears that in the teachers' way of thinking, intelligence and achievement have very much the

same meaning. Consequently, one could reasonably predict that with such an attitude, the teacher might well communicate to the pupil his expectations for the pupil's performance simply by being informed or thinking he is informed, of the pupil's ability level.

The concept "tests" was considerably further from teaching and pupils than from intelligence and achievement, but obviously stood by itself. Likewise, the concept mediocrity was singular in its meaning, received the lowest mean score of all attitudes rated, and was extremely distant in semantic space from all other concepts tested. This finding would seem to lend credence to the notion that the typical middle-class teacher considers a condition of average or mediocre to be a low value, and not desirable in a school setting. It appears, in fact, that the values of the middle-class teacher most probably favor the above-average pupil.

Conclusion and Implications for Future Research

The findings of the present study show that teachers do behave differently toward those pupils whom they consider to be intellectually superior than toward those they consider to be only average in intellectual ability. They make more direct references to the special pupil's being able to achieve; and are more friendly, encouraging, and accepting of the allegedly brighter pupils, expressing such acceptance both verbally and non-verbally. In addition, teachers spend

more time communicating to their "superior" pupils.

Although there is little evidence that teacher expectations affect pupil performance at the secondary school level, a pupil's IQ, achievement, and attitudes might very well be affected if all of his teachers expect him to perform above average. In this experiment, only one-third of the teachers who worked with the pupils during the day expected anything special from them.

The evidence from all studies to date indicates a need for a comprehensive, intensive study of the way in which teachers communicate their prophecies. The dimensions examined in the present experiment were extremely gross. A refinement of the category instrument and improved, more elaborate observation techniques would no doubt prove very useful.

An area of study obviously related to the effectiveness of the teacher's communication of his expectancies is the self-concept of the pupil. To what degree does a teacher's expectations about a pupil's performance affect the pupil's self-concept? At what age does the teacher's expectations diminish so as to become insignificant? What dimensions of pupil self-concept are operative in the interpersonal self-fulfilling prophecy? All these questions need further investigation.

Finally, the entire area of teacher attitudes needs to be researched. Not only teacher attitudes toward pupil a-

bility level, but attitudes toward themselves, toward all facets of school life, and toward various personal characteristics of other people need to be studied.

The demands made upon the schools, the pupils, the teachers, and the experimenters in such research efforts promise to be staggering, but not prohibitive. On the other hand, the information gained about teacher-pupil relationships, the teaching-learning processes, and the general social structure of the typical American educational setting promises to be most rewarding.

CHAPTER V

SUMMARY

The purpose of the present study was to investigate the way in which teachers communicate their expectations about pupil performance to the pupils, and to determine the effects of such teacher expectancies on pupil achievement and attitudes. The basic assumption underlying the study was that a prophecy can be instrumental in its own fulfillment (Merton, 1948). Research by Rosenthal and Jacobson (1968), Flowers (1966), and Pitt (1956) provided the basis for the present experiment.

The following predictions were made: (1) Pupils who are expected by their teachers to achieve well will perform better in the subjects tested than those pupils of like ability of whom nothing special is expected. (2) Pupils who are expected by their teachers to achieve well will develop attitudes which are more favorable toward school, the teacher, self, subjects tested, and fellow students than those of like ability of whom nothing special is expected. (3) Teachers will communicate in a more positive-accepting-supportive manner to the experimental pupils than to the

control pupils. (4) Teachers will communicate in a more negative-rejecting-defensive manner to the control pupils than to the experimental pupils. (5) The effect of teacher expectations as depicted by pupil gains in achievement is greater in the language arts than in mathematics. (6) Teachers spend more time communicating with pupils whom they consider to be "bright" than with pupils of like ability, but whom they consider to be "average." (7) Teachers communicate more often with pupils whom they consider to be "bright" than with pupils of like ability, but whom they consider to be "average." (8) The degree to which a teacher's expectations influence pupil achievement and attitudes is a function of the amount of time the teacher spends communicating with the pupils. (9) The degree to which a teacher's expectations influence pupil achievement and attitudes is a function of the number of occurrences of teacher communication to the pupil. (10) Positive-accepting-supportive pupil communication to teacher increases the magnitude of positive-accepting-supportive teacher communication to pupil.

The sample of subjects was selected from incoming seventh-grade classes and seventh grade faculties in six different junior high schools. Commercial instruments were used to measure IQ and achievement. A category system based on Bale's (1950) Interaction Process Analysis was devised to measure teacher-pupil interaction. Attitude

measures were obtained by using a system of semantic differential scales. Each teacher subject was assigned an experimental group of pupil subjects and a control group of pupil subjects matched for IQ. Pretests of pupil achievement in mathematics and language arts, of IQ, and of attitudes toward school-related concepts were given. During the following nine weeks, four one-hour observations were made of teacher-pupil interaction. At the ninth week, posttests were administered for pupil achievement, IQ, and attitude. In addition, teacher attitudes toward several school-related concepts were measured, and a questionnaire related to the teachers' knowledge of the experiment was administered.

The data were analyzed by means of analysis of variance, trend analysis, and the D statistic. The results were interpreted in terms of the quality of teacher-pupil interaction; the amount of time and occurrences of teacher-pupil interaction; the effect of the experimental condition on pupil achievement and attitudes; the effect of the amount of time and the occurrences of teacher communication on pupil achievement and attitudes; and teacher attitudes toward their profession, their pupils, and other concepts related to achievement and aptitude.

As predicted in hypothesis 3, teachers communicated in a more friendly, encouraging manner to pupils whom they considered to be bright than to those whom they considered to be average. Teachers also spent significantly more time com-

municating with those pupils who were allegedly brighter (hypothesis 6).

Teachers did not appear to communicate more negatively toward the control pupils than toward the experimental pupils as predicted in hypothesis 4. They simply communicated less with those pupils who were not labeled as "intellectually superior."

The prediction made in hypothesis 10 was that the manner of pupil communication would determine the teachers' mode of communication. When the pupils communicated in a positive, accepting manner, the positive nature of teacher communication was enhanced.

Hypothesis 1 received no support. The effect of the experimental condition on pupil achievement was not significant. However, hypothesis 2, predicting that teacher expectations would affect pupil attitudes, was partially supported. Interaction among pupil attitudes was significant, but the difference between language arts and mathematics accounted for most of the variance. In addition, experimental pupils changed their attitudes to a significant degree in favor of language arts as opposed to mathematics. Contrary to hypotheses 8 and 9, neither changes in pupil achievement nor pupil attitudes depended upon the amount of time teachers spent communicating with the pupils or the occurrences of teacher to pupil communication.

The findings of the present study show that teachers

do behave differently toward those pupils whom they consider to be intellectually superior than toward those they consider to be only average in intellectual ability. They make more direct references to the special pupil's being able to achieve; and are more friendly, encouraging, and accepting of the allegedly brighter pupils, expressing such acceptance both verbally and non-verbally. In addition, teachers spend more time communicating to their "superior" pupils.

Although there is little evidence that teacher expectations affect pupil performance at the secondary school level, a pupil's IQ, achievement, and attitudes might very well be affected if all of his teachers expect him to perform above average. In this experiment, only one-third of the teachers who worked with the pupils during the day expected anything special from them.

Suggestions for future research included (1) an intensive and comprehensive study of all relevant dimensions of teacher-pupil communication; (2) the relationship between teacher expectations and pupil self-concept; (3) and the role of teacher attitudes in prophecy communication, and the effect of teacher-pupil interaction upon teacher attitudes.

REFERENCES

Adair, J. G., & Epstein, J. Verbal cues in the mediation of experimenter bias. Paper read at Midwestern Psychological Association, Chicago, May, 1967. Cited by R. Rosenthal & L. Jacobson, Pygmalion in the classroom. New York: Holt, Rinehart, & Winston, 1968. Pp. 28-29.

Allport, G. W. The role of expectancy. In H. Cantril (Ed.), Tensions that cause wars. Urbana, Illinois: University of Illinois, 1950. Pp. 43-78.

Amidon, E. J., & Hunter, E. Improving teaching: the analysis of classroom verbal interaction. New York: Holt, Rinehart, & Winston, 1966.

Amidon, E. J., & Hough, J. B. (Eds.) Interaction analysis: theory, research, and application. Reading, Massachusetts: Addison-Wesley, 1967.

Anderson, H. H., & Brewer, H. M. Studies of teachers' classroom personalities. II. Effects of teachers' dominative and integrative contacts on children's classroom behavior. Applied Psychological Monographs, 1946, 8.

Anderson, H. M. A study of certain criteria of teaching effectiveness. Journal of Experimental Education, 1945, 23, 41-71.

Bales, R. F. Interaction process analysis. Cambridge, Massachusetts: Addison-Wesley, 1950.

Bales, R. F. Conceptual frameworks for analysis of social interaction. Journal of Experimental Education, 1962, 30, 4, 323-324.

Bavelas, A. Personal communication, December 6, 1965. Cited by R. Rosenthal & L. Jacobson, Pygmalion in the classroom. New York: Holt, Rinehart, & Winston, 1968. P. 6.

Beecher, H. K. Pain: one mystery solved. Science, 1966, 151, 840-841.

Bellack, A. A., et al. The language of the classroom: meanings communicated in high school teaching. Unpublished manuscript, Bureau of Publications, Teachers College, Columbia Teachers College, Columbia University, 1963.

Biddle, B. J. The integration of teacher effectiveness research. In B. J. Biddle & W. J. Ellena (Eds.), Contemporary research on teacher effectiveness. New York: Holt, Rinehart, & Winston, 1964. Pp. 1-40.

Biddle, B. J., & Adams, R. S. An analysis of classroom activities. Center for Research in Social Behavior, University of Missouri, 1967.

Bowers, N. D., & Soar, R. S. The influence of teacher personality on classroom interaction. Journal of Experimental Education, 1962, 30, 4, 309-311.

Cahen, L. S. Experimental manipulation of bias in teachers' scoring of subjective tests. Paper read at American Psychological Association, New York, September, 1966. Cited by R. Rosenthal & L. Jacobson, Pygmalion in the classroom. New York: Holt, Rinehart, & Winston, 1968. P. 54.

Chapple, E. D. The interaction chronograph: its evolution and present application. Personnel, 1949, 25, 295-307.

Conn, L. K., Edwards, C. N., Rosenthal, R., & Crowne, D. F. Emotion perception and response to teacher expectancy in elementary school children. Unpublished manuscript, Harvard University, 1967.

Cornell, F. G., Lindvall, C. M., & Saupe, J. L. An exploratory measurement of individualities of schools and classrooms. Urbana, Illinois: Bureau of Educational Research, University of Illinois, 1952.

Crowne, D. P., & Marlowe, D. The approval motive. New York: Wiley, 1964.

Dahlke, A. E. An integrated approach to the analysis of variance. Unpublished manuscript, University of Oklahoma, 1966.

Deutsch, Martin. The disadvantaged child and the learning process. In A. H. Passow (Ed.), Education in depressed areas. New York: Bureau of Publications, Teachers College, Columbia University, 1963. Pp. 163-179.

Deutsch, M, Fishman, J. A., Kogan, L., North, R., & Whiteman, M. Guidelines for testing minority group children. Journal of Social Issues, 1964, 20, 129-145.

Edwards, A. L. Experimental design in psychological research. (Rev. ed.) New York: Holt, Rinehart, & Winston, 1950. Ch. 14, pp. 264-283.

Festinger, L. A theory of cognitive dissonance. New York: Harper & Row, 1957.

Flanders, N. A. Interaction analysis in the classroom: a manual for observers. Unpublished manuscript, University of Michigan, 1960.

Flanders, N. A. Some relationships among teacher influence, pupil attitudes, and achievement. In B. J. Biddle & W. J. Ellena (Eds.), Contemporary research on teacher effectiveness. New York: Holt, Rinehart, & Winston, 1964. Pp. 196-231.

Flowers, C. E. Effects of an arbitrary accelerated group placement on the tested academic achievement of educationally disadvantaged students. Unpublished doctoral dissertation, Teachers College, Columbia University, 1966.

Fode, K. L. The effect of non-visual and non-verbal interaction on experimenter bias. Unpublished master's thesis, University of North Dakota, 1960.

Fode, K. L. The effects of experimenters' and subjects' anxiety and social desirability on experimenter outcome-bias. Unpublished doctoral dissertation, University of North Dakota, 1965.

Friedman, P. A second experiment on interviewer bias. Sociometry, 1942, 5, 378-379.

Gage, N. L. (Ed.) Handbook of research on teaching. Chicago: Rand-McNally, 1963.

Gibson, G. Aptitude tests. Science, 1965, 149, 583.

Harlem Youth Opportunities Unlimited, Inc. Youth in the ghetto. New York: HARYOU, 1964.

Harvey, S. M. Preliminary investigation of the interview. British Journal of Psychology, 1938, 28, 263-287.

Harvey, O. J., & Clapp, W. F. Hope, expectancy, and reactions to the unexpected. Journal of Personality and Social Psychology, 1965, 2, 45-52.

Havighurst, R. J. Education for the great society. The Instructor, September, 1965, 75, 31, 62, 65-66.

Hughes, M. Development of the means for the assessment of the quality of teaching in the elementary schools. Salt Lake City: University of Utah Press, 1959.

Hughes, M. Utah study of the assessment of teaching. In A. A. Bellack (Ed.), Theory and research in teaching. New York: Columbia University, Teachers College, Bureau of Publications, 1963. Pp. 25-36.

Kanfer, F. H., & Karas, S. C. Prior experimenter-subject interaction and verbal conditioning. Psychological Reports, 1959, 5, 345-353.

Kelley, T. L., Madden, R., Gardner, E. F., & Rudman, H. C. Stanford Achievement Test: Advanced Partial Battery. New York: Harcourt, Brace & World, 1964.

Larrabee, L. L., & Kleinsasser, F. The effect of experimenter bias on WISC performance. Unpublished paper. St. Louis, Missouri: Psychological Associates, 1967.

Lewis, W. W., Newell, M. M., & Withall, J. An analysis of classroom patterns of communication. Psychological Reports, 1961, 9, 211-219.

Levy, L. H., & Orr, T. B. The social psychology of Rorschach validity research. Journal of Abnormal and Social Psychology, 1959, 58, 79-83.

Levy, L. H., & Orr, T. B. The social psychology of Rorschach validity research. Journal of Abnormal and Social Psychology, 1964, 69, 572-574.

Lewis, W. W., & Newell, J. M. Analysis of classroom interaction through communication behaviors. Journal of Experimental Education, 1962, 30, 4, 321-322.

Lindguist, E. F. Design and analysis of experiments in psychology and education. Cambridge, Massachusetts: Riverside Press, 1953.

Lindzey, G. A note on interviewer bias. Journal of Applied Psychology, 1951, 35, 182-184.

Loban, W. D. Mastery of conventional usage and grammar. Unpublished report to the U. S. Office of Education, 1964.

McClelland, D. C. The achieving society. Princeton, N. J.: Van Nostrand, 1961.

Marwit, S. J., & Marcia, J. E. Tester-bias and response to projective instruments. Journal of Consulting Psychology, 1967, 31, 253-258.

Masling, J. The influence of situational and interpersonal variables in projective testing. Psychological Bulletin, 1960, 57, 65-85.

Masling, J. Differential indoctrination of examiners and Rorschach responses. Journal of Consulting Psychology, 1965, 29, 198-201.

Medley, D. M., & Mitzel, H. E. A technique for measuring classroom behavior. Journal of Educational Psychology, 1958, 49, 86-92.

Medley, D. M., & Mitzel, H. E. Some behavioral correlates of teacher effectiveness. Journal of Educational Psychology, 1959, 50, 239-246.

Medley, D. M., & Mitzel, H. E. A tentative framework for the study of effective teacher behavior. Journal of Experimental Education, 1962, 30, 4, 317-320.

Merton, R. K. The self-fulfilling prophecy. Antioch Review, 1948, 8, 193-210.

Meux, M., & Smith, O. Logical dimensions of teaching behavior. In B. J. Biddle & W. J. Ellena (Eds.), Contemporary research on teacher effectiveness. New York: Holt, Rinehart, & Winston, 1964. Pp. 127-164.

Mitzel, H. E., & Rabinowitz, W. Assessing social-emotional climate in the classroom by Withall's technique. Psychological Monographs, 1953, 67, 18.

Osgood, C. E. The measurement of meaning. Urbana, Illinois: University of Illinois Press, 1957.

Otis, A. S., & Lennon, R. T. Otis-Lennon Mental Ability Test: Intermediate Level. New York: Harcourt, Brace & World, 1967.

Passow, A. H. (Ed.) Education in depressed areas. New York: Bureau of Publications, Teachers College, Columbia University, 1963.

Pitt, C. C. V. An experimental study of the effects of teachers' knowledge or incorrect knowledge of pupil IQ's on teachers' attitudes and practices and pupils' attitudes and achievement. Unpublished doctoral dissertation, Columbia University, 1956.

Rosenthal, R. On the social psychology of the psychological experiment: the experimenter's hypothesis as unintended determinant of experimental results. American Scientist, 1963, 51, 268-283.

Rosenthal, R. Experimenter effects in behavioral research. New York: Appleton, 1966.

Rosenthal, R., & Fode, K. L. The effect of experimenter bias on the performance of the albino rat. Behavioral Science, 1963, 8, 183-189.

Rosenthal, R., Fode, K. L., Vikan-Kline, L., & Persinger, G. W. Verbal conditioning: Mediator of experimenter expectancy effects? Psychological Reports, 1964, 14, 71-74.

Rosenthal, R., & Halas, E. S. Experimenter effect in the study of invertebrate behavior. Psychological Reports, 1962, 11, 251-256.

Rosenthal, R., & Jacobson, L. Teachers' expectancies: determinants of pupils' IQ gains. Psychological Reports, 1966, 19, 115-118.

Rosenthal, R., & Jacobson, L. Pygmalion in the classroom. New York: Holt, Rinehart, & Winston, 1968.

Rosenthal, R., & Lawson, R. A longitudinal study of the effects of experimenter bias on the operant learning of laboratory rats. Journal of Psychiatric Research, 1964, 2, 61-72.

Scott, W. A. Reliability of content analysis: the case of nominal scale coding. The Public Opinion Quarterly, 1955, 19, 3, 321-325. In N. A. Flanders, Interaction analysis in the classroom: a manual for observers. Unpublished manuscript, University of Michigan, 1960.

Sexton, P. C. Education and income. New York: Viking, 1961.

Shapiro, A. K. A contribution to a history of the placebo effect. Behavior Science, 1960, 5, 109-135.

Sheard, M. H. The influence of doctor's attitude on the patient's response to antidepressant medication. Journal of Nervous and Mental Diseases, 1963, 136, 555-560.

Smith, B. O. A conceptual analysis of instructional behavior. Journal of Teacher Education, 1963, 14, 3, 294-298.

Stanton, F., & Baker, K. H. Interview bias and the recall of incompletely learned materials. Sociometry, 1942, 5, 123-134.

Stevenson, H. W. Social reinforcement of children's behavior. In L. P. Lipsitt & C. C. Spiker (Eds.), Advance in child development and behavior. Vol. 2. New York: Academic, 1965. Pp. 97-126.

Thomas, W. I. Social behavior and personality. E. H. Volkart (Ed.) New York: Social Science Research Council, 1951.

Whyte, W. F. Street corner society. Chicago: University of Chicago Press, 1943.

Withall, J. Development of a technique for the measurement of socio-emotional climate in classrooms. Journal of Experimental Education, 1949, 17, 347-361.

Withall, J. The development of a climate index. Journal of Educational Research, 1951, 45, 93-99.

Wyatt, D. F., & Campbell, D. T. A study of interviewer bias as related to interviewers' expectations and own opinions. International Journal of Opinion and Attitude Research, 1950, 4, 77-83.

APPENDIX A

Definition of Categories

1. Direct Reference (Positive)--Verbal: Any direct verbal reference by either the teacher or the pupil about the pupil's ability level, e. g., "You are too bright to waste your time," etc.
2. Accepting/Supportive--Verbal: All verbalizations, other than direct references to the pupil's ability level, which indicate praise, acceptance, extolling, reassurance, comforting, sympathy, concurrence, acceptance, joking, laughing with, kidding, greeting, e. g., "Yes, that's right," or "I understand just how you feel," or "See you tomorrow," etc.
3. Accepting/Supportive--Non-Verbal: Behaviors that are of the same intent as the Accepting/Supportive verbal category, but are expressed non-verbally, e. g., smiling, waving, nodding approval, admiring, sharing, doing for, laying hand on shoulder approvingly, horseplaying, etc.
4. Task Oriented--Neutral: Gives orientation, information; repeats, clarifies, confirms, asks orientation, information, opinion, or analysis with no tendency toward

positive or negative behavior.

5. Rejecting/Defensive--Non-Verbal: Cool, detached, formal, aloof, unappreciative, inattentive, failing to give assistance, frowning, shaking head or finger in disapproval, ignoring, punishing physically, etc.
6. Rejecting/Defensive--Verbal: Disapproving, defensive or negative tension releasing verbalizations, e. g., disagree, sarcastic, critical, harping, badgering, nagging, teasing, scoffing, extreme arbitrariness, giving warnings or threats, etc.
7. Direct Reference (Negative)--Verbal: Any direct verbal reference by either the teacher or the pupil about the pupil's ability level, e. g., "I'm no genius," or "You may not be the smartest student in the world, but . . .", etc.

APPENDIX B

CSO Score Sheet

Sch_____ Tchr_____ P_E Abs _____ P_C Abs_____ Obs Pd_____

CATEGORY		T-P_E	P_E-T	T-P_C	P_C-T
DR (Pos)	1				
A/S V	2				
A/S N-V	3				
TO (Neut)	4				
R/D N-V	5				
R/D V	6				
DR (Neg)	7				

- COMMENTS -

APPENDIX C

OTIS-LENNON MENTAL ABILITY TEST*

Intermediate Level, Form J

New York: Harcourt, Brace & World

Copyright, 1967

*Results in Appendix I reproduced from Otis-Lennon Mental Ability Test, copyright, 1967, by Harcourt, Brace & World, Inc. Reproduced by special permission.

APPENDIX D

STANFORD ACHIEVEMENT TEST*

Advanced, Partial Battery, Form X

New York: Harcourt, Brace & World

Copyright, 1967

*Results in Appendix I reproduced from Stanford Achievement Tests, copyright, 1964, by Harcourt, Brace & World, Inc. Reproduced by special permission.

APPENDIX E

SEMANTIC DIFFERENTIAL: PUPIL

Instructions

The purpose of this study is to measure the meanings of certain things to various people by having them judge them against a series of descriptive scales. In taking this test, please make your judgments on the basis of what these things mean to you. On each page you will find a different concept to be judged and beneath it a set of scales. You are to rate the concept on each of these scales in order.

Here is how you are to use these scales:

If you feel that the concept at the top of the page is very closely related to one end of the scale, you should place your check-mark as follows:

fair __X__:____:____:____:____:____:____ unfair

or

fair ____:____:____:____:____:____:__X__ unfair

If you feel that the concept is quite closely related to one or the other end of the scale (but not extremely), you should place your check-mark as follows:

strong ____:__X__:____:____:____:____:____ weak

or

strong ____:____:____:____:____:____:__X__ weak

If the concept seems only slightly related to one side as opposed to the other side (but is not really neutral), then you should check as follows:

active _____:_____:__X__:_____:_____:_____:_____ passive

or

active _____:_____:_____:_____:__X__:_____:_____ passive

The direction toward which you check, of course, depends upon which of the two ends of the scale seem most characteristic of the thing you're judging. If you consider the concept to be neutral on the scale, both sides of the scale equally associated with the concept, or if the scale is completely irrelevant, unrelated to the concept, then you should place your check-mark in the middle space:

safe _____:_____:_____:__X__:_____:_____:_____ dangerous

IMPORTANT: (1) Place your check-marks in the middle of the spaces, not on the boundaries:

THIS			NOT THIS		
_____:	__X__	:_____	:__X__	:_____	

(2) Be sure you check every scale for every concept--do not omit any.
(3) Never put more than one check-mark on a single scale.

Sometimes you may feel as though you've had the same item before on the test. This will not be the case; so do not look back and forth through the items. Do not try to remember how you checked similar items earlier in the test. Make each item a separate and independent judgment. Work at a fairly high speed through this test. Do not worry or puzzle over individual items. It is your first impressions, the immediate "feelings" about the items, that we want. On the other hand, please do not be careless, because we want your true impressions.

CONCEPTS AND SCALES

TEACHER

good	_____:_____:_____:_____:_____:_____:_____	bad
fair	_____:_____:_____:_____:_____:_____:_____	unfair
unpleasant	_____:_____:_____:_____:_____:_____:_____	pleasant
high	_____:_____:_____:_____:_____:_____:_____	low
dull	_____:_____:_____:_____:_____:_____:_____	smart
tense	_____:_____:_____:_____:_____:_____:_____	relaxed
valuable	_____:_____:_____:_____:_____:_____:_____	worthless

SCHOOL

high	_____:_____:_____:_____:_____:_____:_____	low
dirty	_____:_____:_____:_____:_____:_____:_____	clean
awful	_____:_____:_____:_____:_____:_____:_____	nice
fair	_____:_____:_____:_____:_____:_____:_____	unfair
boring	_____:_____:_____:_____:_____:_____:_____	fun
worthless	_____:_____:_____:_____:_____:_____:_____	valuable
easy	_____:_____:_____:_____:_____:_____:_____	hard

ENGLISH

unfair	_____:_____:_____:_____:_____:_____:_____	fair
high	_____:_____:_____:_____:_____:_____:_____	low
valuable	_____:_____:_____:_____:_____:_____:_____	worthless
fun	_____:_____:_____:_____:_____:_____:_____	boring
hard	_____:_____:_____:_____:_____:_____:_____	easy
bad	_____:_____:_____:_____:_____:_____:_____	good
clear	_____:_____:_____:_____:_____:_____:_____	hazy

MATH

boring	_____:_____:_____:_____:_____:_____:_____	fun
hard	_____:_____:_____:_____:_____:_____:_____	easy
valuable	_____:_____:_____:_____:_____:_____:_____	worthless
low	_____:_____:_____:_____:_____:_____:_____	high
fair	_____:_____:_____:_____:_____:_____:_____	unfair
hazy	_____:_____:_____:_____:_____:_____:_____	clear
unpleasant	_____:_____:_____:_____:_____:_____:_____	pleasant

FELLOW STUDENT

clean _____:_____:_____:_____:_____:_____:_____ dirty

smart _____:_____:_____:_____:_____:_____:_____ dull

awful _____:_____:_____:_____:_____:_____:_____ nice

fair _____:_____:_____:_____:_____:_____:_____ unfair

dishonest _____:_____:_____:_____:_____:_____:_____ honest

bad _____:_____:_____:_____:_____:_____:_____ good

relaxed _____:_____:_____:_____:_____:_____:_____ tense

ME

dishonest _____:_____:_____:_____:_____:_____:_____ honest

dull _____:_____:_____:_____:_____:_____:_____ smart

valuable _____:_____:_____:_____:_____:_____:_____ worthless

bad _____:_____:_____:_____:_____:_____:_____ good

high _____:_____:_____:_____:_____:_____:_____ low

calm _____:_____:_____:_____:_____:_____:_____ agitated

cruel _____:_____:_____:_____:_____:_____:_____ kind

APPENDIX F

SEMANTIC DIFFERENTIAL: TEACHER

Instructions

The purpose of this study is to measure the meanings of certain things to various people by having them judge them against a series of descriptive scales. In taking this test, please make your judgments on the basis of what these things mean to you. On each page you will find a different concept to be judged and beneath it a set of scales. You are to rate the concept on each of these scales in order.

Here is how you are to use these scales:

If you feel that the concept at the top of the page is very closely related to one end of the scale, you should place your check-mark as follows:

fair __X__:_____:_____:_____:_____:_____:_____ unfair

or

fair _____:_____:_____:_____:_____:_____:__X__ unfair

If you feel that the concept is quite closely related to one or the other end of the scale (but not extremely), you should place your check-mark as follows:

strong _____:__X__:_____:_____:_____:_____:_____ weak

or

strong _____:_____:_____:_____:_____:__X__:_____ weak

If the concept seems only slightly related to one side as opposed to the other side (but is not really neutral), then you should check as follows:

active _____:_____:__X__:_____:_____:_____:_____ passive

or

active _____:_____:_____:_____:__X__:_____:_____ passive

The direction toward which you check, of course, depends upon which of the two ends of the scale seem most characteristic of the thing you're judging. If you consider the concept to be neutral on the scale, both sides of the scale equally associated with the concept, or if the scale is completely irrelevant, unrelated to the concept, then you should place your check-mark in the middle space:

safe _____:_____:_____:__X__:_____:_____:_____

IMPORTANT: (1) Place your check-marks in the middle of the spaces, not on the boundaries:

THIS			NOT THIS		
_____:	__X__	:_____	:X_____		:_____

(2) Be sure you check every scale for every concept--do not omit any.
(3) Never put more than one check-mark on a single scale.

Sometimes you may feel as though you've had the same item before on the test. This will not be the case; so do not look back and forth through the items. Do not try to remember how you checked similar items earlier in the test. Make each item a separate and independent judgment. Work at fairly high speed through the test. Do not worry or puzzle over individual items. It is your first impressions, the immediate "feelings" about the items, that we want. On the other hand, please do not be careless, because we want your true impressions.

CONCEPTS AND SCALES

TEACHING

fair	_____:_____:_____:_____:_____:_____:_____	unfair
good	_____:_____:_____:_____:_____:_____:_____	bad
weak	_____:_____:_____:_____:_____:_____:_____	strong
happy	_____:_____:_____:_____:_____:_____:_____	sad
dishonest	_____:_____:_____:_____:_____:_____:_____	honest
easy	_____:_____:_____:_____:_____:_____:_____	hard
unpleasant	_____:_____:_____:_____:_____:_____:_____	pleasant

INTELLIGENCE

fair	_____:_____:_____:_____:_____:_____:_____	unfair
good	_____:_____:_____:_____:_____:_____:_____	bad
weak	_____:_____:_____:_____:_____:_____:_____	strong
happy	_____:_____:_____:_____:_____:_____:_____	sad
dishonest	_____:_____:_____:_____:_____:_____:_____	honest
easy	_____:_____:_____:_____:_____:_____:_____	hard
unpleasant	_____:_____:_____:_____:_____:_____:_____	pleasant

TESTS

fair	_____:_____:_____:_____:_____:_____:_____	unfair
good	_____:_____:_____:_____:_____:_____:_____	bad
weak	_____:_____:_____:_____:_____:_____:_____	strong
happy	_____:_____:_____:_____:_____:_____:_____	sad
dishonest	_____:_____:_____:_____:_____:_____:_____	honest
easy	_____:_____:_____:_____:_____:_____:_____	hard
unpleasant	_____:_____:_____:_____:_____:_____:_____	pleasant

PUPILS

fair	_____:_____:_____:_____:_____:_____:_____	unfair
good	_____:_____:_____:_____:_____:_____:_____	bad
weak	_____:_____:_____:_____:_____:_____:_____	strong
happy	_____:_____:_____:_____:_____:_____:_____	sad
dishonest	_____:_____:_____:_____:_____:_____:_____	honest
easy	_____:_____:_____:_____:_____:_____:_____	hard
unpleasant	_____:_____:_____:_____:_____:_____:_____	pleasant

ACHIEVEMENT

fair	_____:_____:_____:_____:_____:_____:_____	unfair
good	_____:_____:_____:_____:_____:_____:_____	bad
weak	_____:_____:_____:_____:_____:_____:_____	strong
happy	_____:_____:_____:_____:_____:_____:_____	sad
dishonest	_____:_____:_____:_____:_____:_____:_____	honest
easy	_____:_____:_____:_____:_____:_____:_____	hard
unpleasant	_____:_____:_____:_____:_____:_____:_____	pleasant

MEDIOCRITY

fair	_____:_____:_____:_____:_____:_____:_____	unfair
good	_____:_____:_____:_____:_____:_____:_____	bad
weak	_____:_____:_____:_____:_____:_____:_____	strong
happy	_____:_____:_____:_____:_____:_____:_____	sad
dishonest	_____:_____:_____:_____:_____:_____:_____	honest
easy	_____:_____:_____:_____:_____:_____:_____	hard
unpleasant	_____:_____:_____:_____:_____:_____:_____	pleasant

APPENDIX G

KNOWLEDGE OF EXPERIMENT

1. Do you suspect that the real purpose of this experiment was different than you were originally led to believe?

2. If the answer to question number one is "yes," what do you think really is the purpose of the experiment? (If your answer to question number one was "no," just leave this question blank.)

3. When, if ever, did you begin to suspect that the purpose of the experiment might be different than originally stated?

4. What led you to suspect that the purpose of the experiment was different than that which was stated?

APPENDIX H

COMMUNIQUE TO TEACHERS

To: (Name of Teacher)

From: S. W. Kester, Ass't. Prof. of Educ., O. B. U.

SUBJECT: STUDY OF INTELLECTUALLY SUPERIOR PUPILS

Dear (Mr./Mrs./Miss) _____:

As a part of a research project related to the functioning of intellectually superior children, several exceptionally bright pupils (IQ greater than 120, Otis-Lennon MAT) have been identified in this school. The following have been assigned to your _____ hour class, which is one of the classes to be used in the study:

(Names of Pupils)

It is our purpose to observe and record the behavior of these pupils in an "average" classroom setting. An observer-experimenter will be contacting you to ask your cooperation. We realize that your participation will be somewhat of an inconvenience to you and an intrusion in your

class, but we will do everything possible to make our presence unimposing!

The following procedures would help us immeasurably:

1. To avoid the "guinea pig" effect among the pupils, perhaps you might announce the observer as, ". . . a student from O. B. U. who is trying to learn about what takes place in the classroom. He/she will be with us four or five times this nine weeks; so we will just go on about our business," or something to this effect.
2. Because "guinea pig" effects are picked up from many sources, we also ask that you not discuss the experiment with anyone--pupils, fellow teachers, counselors, administrators--until after the first quarter has ended and all data have been collected. You will then be provided with a full report of the study.
3. If you would, please, allow the observer to seat himself so that he can see the faces of all the pupils in the room.
4. It would also be extremely helpful if you would assist the observer in filling out a seating chart and inform him of any changes that you make between observations.

We realize that without the cooperation of the classroom teacher, very little information useful to education can really

be gathered. We want you to know that we appreciate your giving of your time and effort.

APPENDIX I

IBM CARD FORMATS AND PRINTOUTS OF DATA

Column Key	Pretest and Posttest Data for Pupil Achievement and IQ

Pretest Data Only for Pupil Attitudes; Card 01

1-3.	Pupils' subject identification numbers
4.	Sex (1=male; 2=female)
5-6.	Teacher identification number
7.	Group (1=experimental; 2=control)
8-10.	Pretest scores in mathematics
11-13.	Posttest scores in mathematics
14.	Direction of change in score (2=up; 1=down; 0=no change)
15-16.	Difference scores in mathematics
17-19.	Pretest scores in language arts
20-22.	Posttest scores in language arts
23.	Direction of change in score (2=up; 1=down; 0=no change)
24-25.	Difference scores in language arts
26-28.	Pretest scores in IQ
29-31.	Posttest scores in IQ
32.	Direction of change in score (2=up; 1=down; 0=no change)
33-34.	Difference scores in IQ
35-41.	Pretest scores in pupil attitude toward teacher
42-48.	Pretest scores in pupil attitude toward self
49-55.	Pretest scores in pupil attitude toward school
56-62.	Pretest scores in pupil attitude toward mathematics
63-69.	Pretest scores in pupil attitude toward English
70-76.	Pretest scores in pupil attitude toward fellow student
77-78.	Card identification number

Column-Key	Posttest Data for Pupil Attitudes; Card 02
1-34.	Information repeated from Card 01
35-41.	Posttest scores in pupil attitude toward teacher
42-48.	Posttest scores in pupil attitude toward self
49-55.	Posttest scores in pupil attitude toward school
56-62.	Posttest scores in pupil attitude toward mathematics
63-69.	Posttest scores in pupil attitude toward English
70-76.	Posttest scores in pupil attitude toward fellow student
77-78.	Card identification number

Difference Score Data for Pupil Attitudes; Card 05

1-34.	Information repeated from Card 01
35.	Direction of change in score (blank=up; 1=down; 0=no change)
36-37.	Difference score for pupil attitude toward teacher
38.	Direction score for pupil attitude toward teacher
39-40.	Difference score for pupil attitude toward self
41.	Direction score for pupil attitude toward self
42-43.	Difference score for pupil attitude toward school
44.	Direction score for pupil attitude toward school
45-46.	Difference score for pupil attitude toward mathematics.
47.	Direction score for pupil attitude toward mathematics
48-49.	Difference score for pupil attitude toward English
50.	Direction score for pupil attitude toward English
51-52.	Difference score for pupil attitude toward fellow student
53-54.	Card identification number
55-78.	Blank

Teacher Attitude Scores; Card 06

1-2.	Teacher subject identification number
3.	Sex (1=male; 2=female)
4.	School identification number
5-11.	Scores for teacher attitude toward teaching
12-18.	Scores for teacher attitude toward pupils
19-25.	Scores for teacher attitude toward tests
26-32.	Scores for teacher attitude toward achievement
33-39.	Scores for teacher attitude toward intelligence
40-46.	Scores for teacher attitude toward mediocrity
47-48.	Total score for teacher attitude toward teaching
49-50.	Total score for teacher attitude toward pupils

Column Key

51-52. Total score for teacher attitude toward tests
53-54. Total score for teacher attitude toward achievement
55-56. Total score for teacher attitude toward intelligence
57-58. Total score for teacher attitude toward mediocrity
59-60. Card identification number
61-78. Blank

Observations 1-4 of Communication of T-P_E and P_E-T in Both Frequency and Amount of Time; Cards 07, 09, 11, 13

1-2. Teacher subject identification number
3. Sex (1=male; 2=female)
4-17. T-P_E (frequency); 2 columns per category, beginning with columns 4-5 as category one
18-31. P_E-T (frequency); 2 columns per category, beginning with columns 18-19 as category one
32-45. T-P_E (time); 2 columns per category, beginning with columns 32-33 as category one
46-59. P_E-T (time); 2 columns per category, beginning with columns 46-47 as category one
60. Group (1=experimental; 2=control)
61. Observation number
62-63. Card identification number
64-78. Blank

Observations 1-4 of Communication of T-P_C and P_C-T in Both Frequency and Amount of Time; Cards 08, 10, 12, 14

1-2. Teacher subject identification number
3. Sex (1=male; 2=female)
4-17. T-P_C (frequency); 2 columns per category, beginning with columns 4-5 as category one
18-31. P_C-T (frequency); 2 columns per category, beginning with columns 18-19 as category one
32-45. T-P_C (time); 2 columns per category, beginning with columns 32-33 as category one
46-59. P_C-T (time); 2 columns per category, beginning with columns 46-47 as category one
60. Group
61. Observation number
62-63. Card identification number
64-78. Blank