DOCUMENT RESUME

ED 090 236 95 SP 007 982

AUTHOR Wilkins, William E.
TITLE Teacher Expectations and Classroom Behaviors.
INSTITUTION State Univ. of New York, Brockport. Coll. at Brockport.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE [72]
GRANT OEG-IB-050
NOTE 33p.

EDRS PRICE MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS *Academic Achievement; *Educational Research; *Student Ability; Student Improvement; *Student Motivation; Student Teacher Relationship; *Teacher Attitudes; Teacher Behavior

ABSTRACT

To test the effect of teacher expectation on pupil achievement, subjects in 24 classrooms, grades 1-6, a) completed the Metropolitan Analysis of Learning Potential and the Metropolitan Achievement Test; b) completed a questionnaire regarding perceptions of teachers' differential expectations and treatment of students; and c) were ranked by their teachers according to their expected achievement growth during the year. Observers made regulars visits to the classrooms to measure pupil-teacher interactions. The five experimental hypotheses tested were a) experimental subjects will perform better on the ability posttest than the control subjects; b) experimental subjects will score higher on the achievement posttest than the control subjects; c) experimental subjects in the middle ability range will show greater improvement than the control subjects in the same ability range; d) experimental subjects will receive significantly different instruction than the control subjects as measured by independent observers; and e) pupils will report significantly different treatment of control and experimental subjects. Analysis of the collected data shows no significant difference between the control and experimental groups on ability and achievement posttests and suggests rejection of hypotheses "a," "b," "c," and e." Although there was a difference in the instruction received by different ability groups, there was no difference between that received by control and experimental subjects in the same ability grouping, and hypothesis "d" was rejected as untestable. (HMD)

TEACHER EXPECTATIONS AND CLASSROOM BEHAVIORS

William E. Wilkins

Department of Psychology

SUC at Brockport

SPONSOR

Division C Section C-2

The research reported herein was supported, in part, by USOE Grant IB-050.

Introduction:

The publication of Pygmalion in the Classroom (Rosenthal and Jacobson, 1968) has led to a flurry of research and discussion on teacher expectation and their effects on the academic and social behaviors of students. Critics of the Rosenthal research on experimenter bias and teacher expectations have suggested that the findings reported by Rosenthal and his co-workers may be generally untrustworthy. Conceptual difficulties and methodological errors of the kinds suggested by Barber and Silver (1968), Thorndike (1968) and Elashoff and Snow (1971), among others, indicate that the wide publication and use of the Rosenthal and Jacobson findings is not justified without further documentation.

At least ten other attempts to raise IQ scores through expectancy effects have been reported (Conn, et al., 1968; Evans and Rosenthal, 1969; Clairborn, 1969; Jose and Cody, 1971; Fleming and Anttonen, 1971; Kester, 1971; Goldsmith and Fry, 1970; Anderson and Rosenthal, 1968; and Pellegrini and Hicks, 1972). In no case did the findings support the Oak School Experiment.

On the other hand, the research using achievement gains as dependent variables suggests that teacher expectancies probably have some effect on student achievement under certain naturally occurring conditions or under strong experimental manipulation. Rist's findings (1970) suggest that teacher expectancy affects achievement insofar as the teacher refuses to interact with and teach certain students because of culturally established expectations. Meichenbaum, Bowers, and Ross (1969) demonstrated the self-fulfilling effect using fourteen female adolescent juvenile offenders. In this experiment, objective and subjective exams

prepared by the teachers yielded the dependent variables. Schrank (1968, 1970) reported two successful expectancy studies using college freshmen. Final test scores on a mathematics test served as the criterion measures in both experiments. Palardy (1969) demonstrated that teachers' expectations concerning the probability that boys and girls will learn to read was related to the reading achievement test scores of their students. Other studies have shown that expectations relate to the number and quality of teacher-pupil interaction (e.g. Rothbart, et al., 1971; Rubovits and Maehr, 1970; Good, 1970; and Brophy and Good, 1970).

In sum, the literature suggests that teacher expectations may have an effect on some of the behaviors and performances of their students. However, the same literature suggests many obvious contradictions which warrant further study. The present study was designed to investigate the relationships between teacher expectations and student IQ, student achievement, and the number and quality of pupil-teacher interactions. Three general problems were considered in generating testable hypotheses: 1) Does experimental manipulation of teachers' expectations lead to increases in pupils' ability and achievement scores?; 2) If any score increases do occur, are they a function of the self-fulfiling prophecy or simply the result of different instructional programs?; and, 3) If any increases in scores do occur, and if they are the result of the self-fulfilling prophecy, can we identify some of the processes involved?

Ample evidence has been presented to suggest that the results reported by Rosenthal and Jacobson (1968) may well be spurious. It therefore seems clear that a methodologically sound replication is warranted. Since a major portion of this experiment and the entire Rosenthal version

of teacher expectancy hinges first on increases in criterion measures as a result of experimental manipulation of expectancy, clear and unimpeachable increases in achievement and ability scores must be demonstrated. A second problem relates to the possibility that the teacher may engage in different modes of instruction for these pupils who have been specifically classified. That is, if teachers expect a particular pupil to do especially well, he may place him in an instructional group which has better pupils, better instructional materials, and possibly different modes of instruction. Pupils who had not been marked as deserving special consideration may not be given this attention.

In short, the gains in IQ and achievement which Rosenthal and Jacobson interpreted as expectancy effect caused by some subtle communication system may well have been the result of placing the experimental students into one instructional group rather than into another. This rationale suggests that, if there is a subtle communication system involved in the self-fulfilling prophecy, it may not be as important as the original instructional placement of the students by the teacher. The interpretation posited here may help explain why teachers could recall or recognize the names of only a few of the "bloomers" in the Oak School Experiment (Rosenthal and Jacobson, 1968).

Furthermore, Rosenthal and Jacobson report that experimental Ss in the middle range of ability made greater gains in reading achievement than either the low ability group or the high ability group. Assuming that original instructional placement is more important than the subtle communication system, this finding suggests that teachers more readily believe that they could mistakenly place an average pupil in the wrong group

but not a very competent pupil or a very incompetent pupil.

The third problem to be investigated is under the rubric of the process of the self-fulfilling prophecy and concerns the pupils' and the teachers' perceptions of expectancy. That is, are the expectancies conscious on the part of those involved? Specifically, (1) do the teachers have the expectations during and at the end of the treatment period; (2) do the matched pupils differ from their controls in their perceptions of what the teachers expect for them and for their matched counterparts; and (3) do the other pupils in the class recognize a different communication system between the teacher and the experimental subjects and the teacher and the matched control subjects? In other words, pupils must perceive that the teacher expects different achievement and behavior among them before a pupil can make a response contingent on the teachers' expectations. Therefore, teachers' expectations must be revealed in some form so that pupils can perceive differential treatment.

A review of previous research did not yield information directly concerned with pupils' abilities to perceive teacher expectations and differential treatment. However, indirect support can be taken from Flanagan and Havumaki (1960) who demonstrated that tenth grade pupils praised more often by the experimenter were chosen more frequently on a sociogram. Beck (1964) reported the development of a pupil perception of teacher behavior questionnaire, thereby demonstrating that intermediate level elementary school children could, and do, perceive at least eleven specific teacher characteristics. Lippitt and Gold (1959), and Fox, Lippitt and Schmuck (1964) found that elementary school children perceived seventeen specific characteristics of their peers and that students uti-

lized these perceptions in conceptualizing the power structure of the classroom. Furthermore, these perceptions remained consistent throughout the school year.

In sum, it appears that elementary school children make accurate perceptions of subtle affective and cognitive behaviors of their teachers and their peers. Furthermore, it seems reasonable that pupils can perceive differential expectations of the teacher and that these perceptions may be a vital link in whatever communication system may be involved in the self-fulfilling prophecy process.

Within the framework of these general problems, five specific hypotheses were tested:

1. Experimental Ss will show significantly better performance on the ability posttest than the matched control Ss.
2. Experimental Ss will show significantly better performance on the achieveme t posttest than the matched control Ss.
3. When compared with their matched control subjects, experimental Ss in the middle ranges of ability will make significantly greater improvement in achievement and ability than either the high ability experiemntal Ss or the low ability experimental Ss.
4. Experimental Ss will receive significantly different instruction from that of the matched control Ss as measured by group classification and ratings by independent observers.
5. Pupils will report, on a questionnaire measuring pupils' perceptions of teacher behavior, significantly different pupil-teacher interactions between experimental Ss and the teacher as compared with the interactions of matched control Ss and the teacher.

Procedure:

1. Selection of Subjects and Experimental Design

All students in twenty-four classrooms, grades 1-6 (four classrooms at each grade level) were administered the Metropolitan Analysis of Learning Potential, and the Metropolitan Achievement Test.[1] Students were rank orderdd according to their achievement test scores and their ability test scores. Classroom teachers were asked to rank their students according to how much achievement growth they expected from them during the year. Three matched pairs were then selected from each classroom, one pair to represent each third of the distributions. That is, one experimental pair was selected if each member has an achievement score in the top third of the achievement distribution, if each had a potential score in the top third of the ability distribution, and if each had been ranked by his teacher in the top third of her achievement growth prediction distribution, etc. One further qualification for sampling was that the pairs were matched by sex. Furthermore, there was an equal number of males and females at each grade level, but not for each class in the sample. After the matched pairs had been selected, one member of each pair was randomly assigned to the experimental group with the other member becoming part of the control group. Selection of Ss and implementation of she study was completed within the first four weeks of the fall term. Thus, there were twelve experimental Ss and twelve control Ss at each of the six grade levels, or a total of seventy-two experimental subjects matched with seventy-two control subjects. Classrooms were selected from two school districts on the basis of size and geographical location. Prior to the experiment, a questionnaire to determine pupils'

perceptions of teachers' differential expectations and treatment of students had been devised and tested for feasibility in classrooms not to be used in the study. The instrument was administered to all students in the twenty-four participating classrooms at the beginning of the school year. Trained observers made regular visits to selected classrooms to use the observation scale devised by Rist (1970) as an added measure of pupil-teacher interactions.

It has been hypothesized that instructional groupings of students may be more important than the communication system between the students and teacher. Therefore, a documentation of the grouping of experimental and control subjects was made.

2. Inducing Teacher Expectancy

It was noted earlier that each classroom teacher was asked to rank his students according to how much he expected them to grow in achievement. After the experimental subjects were selected, the investigator had a conference with each teacher for the supposed purpose of discussing the differences between his ranking and the test scores. Because of the wide spread popularization of the Rosenthal findings, teachers were told that further research into the problems of achievement testing was being done. Using this as a ploy, the investigator pointed out to the teacher, during the conference, that the scores obtained on the two tests indicated that he had seriously underestimated the achievement potential of several of his pupils. Further, he was informed that we would like his cooperation in trying to validate our test scores by classroom

observation; the investigator then informed him that three pupils in his classroom, as well as three pupils in each of several other classrooms, had been singled out for class study. He was then given the names of the three experimental subjects and an expectancy inducement. That is, it was pointed out that while he may have ranked subject one as likely to grow, our tests indicated that this student should grow or improve much more rapidly than even he had first anticipated. He was then told that subjects two and three representing the middle third and the bottom third of his ranking, ought to improve much more than he expected. In short, he was strongly induced to believe that these three students were going to make exceptional gains during the course of the year.

3. Observation and Post-Testing

Taachers were informed that observers would make regular visits to his classroom to see if the responses and school work which the experimental Ss produced in the classroom were sonsistent with responses that they made on the tests. It was also aniticpated that the regular visits of these observers would serve as a reminder to the teacher and therby reinforce the expectancy.

During the entire school year, one of three observers made weekly observations of approximately forty-five minutes each in each of the twenty-four classrooms. Observers used the nine item teacher-pupil interaction scale developed by Rist (1970). Observers also noted changes in grouping, and other informal observations about the general atmosphere and operation of the classroom.

At the end of the school year, the seventy-two experimental Ss and the seventy-two control Ss were re-administered the Metropolitan Achievement Test and the Metropolitan Analysis of Learning Potential. The pupil perception of teacher behavior questionnaire, developed by the investigator, was re-administered to all of the children in all twenty-four classrooms. After all posttesting was completed, the teachers were informed of the actual purpose of the experiment.

Results:

The effects of the experimental manipulation of teacher expectations was tested on eight dependent variables. These criterion measures covered a variety of academic and social factors. The mean scores and standard deviations for each measure are presented in the following tables. Note that these descriptive statistics are computed by ability level within grade level for the control Ss and the experimental Ss for each of the eight dependent variables.

---------- INSERT TABLES 1 to 9 ABOUT HERE ----------

Careful inspection of these Tables suggests that significant differences between experimental Ss and control Ss are not likely except on the Observed Interaction variable.

Differences on all variables were tested for significance by analysis of variance. Of particular concern for each dependent variable is the comparison between experimental and control subjects and interactions involving the treatment factor. Matched pairs within grade-ability level-treatment (df=72) served as the

error term for all comparisons. Since exact matching of achievement test scores was not always possible because of other matching criteria, the pretest scores on the Metropolitan Achievement Test and the pretest scores on the Analysis of Learning Potential Test were treated as covariates in all analyses. Two separate analyses were done. The first encompassed grades 1-4 inclusive and the second encompassed grades 3-6 inclusive. Of the available computer program options, it was concluded that retaining the two covariates was more valuable than analyzing all six grades simultaneously without controlling for pretest differences. Therefore, achievement gains served as the first dependent variable and ability gains served as the second criterion measure with the pretest scores on each measure being treated as a covariate.

The third dependent variable tested was termed positive interaction. These data were derived from observations of pupil-teacher interactions. The observation scheme used was developed and reported by Rist (1970). The clearness of the category parameters, together with pre-observation discussions, led to consistent recording of interactions. Three kinds of positive interactions--verbal, non-verbal, and physical--were recorded and combined into a positive interaction score for each control subject and each experimental subject.

The fourth dependent variable, neutral interaction, was derived in the same manner as the third except that interactions defined as neutral were recorded. Variable number five, negative

interaction, parallels variables three and four except that negative interactions, including verbal, non-verbal, and physical were recorded. The sixth dependent variable was the combined interaction score for each Ss. All interactions, from all three modes of interaction (physical, non-verbal, and verbal), were added into a single score.

The final two variables, sociometric-pre and sociometric-post, were used in testing Hypothesis 5. The sociometric instrument used to generate this data was developed by the writer. The results of a factor analysis (principal component method) of the twenty-three item scale yielded a principal component comprised of thirteen items. The scores used in the data analysis are the sum of the number of times each control subject and each experimental subject was chosen by anyone in his classroom on the thirteen items which clustered on the principal component.

As suggested by the tables of means and standard deviations (presented above) the analysis of variance resulted in a small number of significant F ratios. In fact, of the 92 F ratios computed, only 12 reached acceptable levels of significance. Of the 12 significant F-ratios, three, regarding over-all achievement gains for the entire sample were relatively uninteresting. These differences, on Variables 1 and 2 indicate more variance between the pretest scores and the posttest scores than within group variance. This difference was found for Grades 1-4 on both the achievement measure and the ability measure ($F = 21.36$, $p < .001$ and $F = 18.45$, $p < .001$ with 3, 10 df respectively). For the

Grades 3-6 analysis, the difference was found only for ability scores ($F = 6.03$, $p < .013$ with 3 10 df). These findings probably suggest the obvious; namely, that all Ss made significant acheivement gains over the course of a year.

Analysis of data by ability level within grade level yielded only one significant F-ratio. This F indicated a difference in the number of pupil-teacher interactions for experimental and control groups combined for grades 1-4 ($F = 3.22$, $p < .020$, 6, 22 df). Inspection of the mean scores presented previously reveals that, for grade 1 Ss, high ability students had more interactions than did their low ability peers. This finding is in agreement with other findings reported in the literature (e.g., Good and Brophy, 1970). For grade 2 Ss, findings indicated that average ability students and low ability students had significantly more interactions with their teachers than did high ability students. In grade 3, low ability students interacted more with their teachers than either average ability students or high ability students. These findings are incongruent with others previously reported. One possible explanation involves differences in instructional methods used with different ability groups. That is, very often teachers would assign high ability students independent projects and then leave them to their own initiative while she spent more time in more directed activities with average and low ability students. An inspection of scores for grade 4 Ss revealed no significant differences in the number of pupil-teacher interactions for the various ability levels.

When the data were analyzed by ability level without regard to grade level, two significant F-ratios were found. Both of these differences involved the observed interactions between pupils and teachers. For Grades 1-4, differences were found in the total number of pupil-teacher interactions (F = 4.75, p. less than .019, 2, 22 df). Inspection of raw scores indicates that low ability students had the most interactions with their teachers, followed by the average ability students, with high ability students, having the fewest interactions with their teachers.

For Grades 3-6, significant differences involved the positive interaction variable (F = 4.48, p. less than .023, 2 22 df). Further analysis revealed that tw ability students had the largest number of positive interactions with their teachers, average ability students had the smallest number of positive interactions with their teacher; and high ability students had more positive interactions with their teacher than did the average ability student but fewer than the low ability students.

The explanation for these findings is conjecture. The independent study hypothesis offered above might also apply to these results. Another possible alternative is that the absolute number of interactions is an artifact of the observatinn time. While matched pairs were observed for an equal amount of time, it is conceivable that observation time across grade level and within ability group, but without regard to the matched pair criterion, may not have been equal. Thus, low ability students may have been involved with the teacher for a larger share of the observation time than either of the two groups. This possibility is partiularly appropriate to the lower grades where a large amount of instructional

time is spent in small group instruction.

Testing for the treatment effect, experimental manipulation of teacher expectancy, between experimental Ss and control Ss revealed no significant F-ratios. The failure to find any differences involving the treatment factor on any of the various academic and social variables suggests rejection of Hypothesis 1, Hypothesis 2, and Hypothesis 5. These results would be consistent with the majority of studies discussed previously.

When treatment effects were tested without regard to grade level, five significant F-ratios were obtained. These positive findings also inovlved the observed pupil-teacher interaction . Inspection of the raw data indicated that experimental subjects had more positive and neutral interactions with their teachers than the control subjects in Grades 1-4 when grade level and ability level were not considered ($F = 5.30$, p. less than .044; $F = 13.28$, p. less than .005, 1, 10 df respectively). An examination of the raw scores relevant to Grades 3-6 suggests that experimental Ss had more total interactions ($F = 9.07$, p. less than .013, 1, 10 df) as well as more neutral interactions ($F = 6.75$, p. less than .028, 1, 10 df) and more positive interactions ($F = 5.66$, p. less than .039, 1, 10 df) with their teachers when compared with the control Ss. While these findings seem to support the general thesis of teacher expectancy effects, an accurate interpretation of these findings is unlikely.

Since the design utilized in this experiment was specifically developed to test within grade level differences, differences across grade levels cannot be accounted for. It is possible that the differences across grade levels is a function of between grade variance rather than between experimental Ss and control Ss variance. In any event, these findings

are at best weakly supportive of Hypothesis 5.

When treatment effects were tested within grade level and by ability, no significant differences were obtained between the matched pairs. The failure to find interaction effects rules out the possibility of accepting any specific hypotheses. The lack of significant interaction effects for the treatment factor, combined with the failure to find any significant grade level-treatment factor, combined with the failure to find any significant grade level-treatment interaction effects forces a rejection of all hypotheses except number 4. Since Hypothesis 4 concerns differences in instructional methods rather than differences on any of the eight dependent variables, it has yet to be considered. Before considering data relevant to this hypothesis, some closure on the analysis of variance is necessary. A single significant F-ratio was obtained when the treatment effects for different ability levels was tested without regard for grade level. Again, the observed pupil-teacher interaction variable is involved ($F = 9.41$, p. less than .001, 2, 22df). Analysis of raw scores indicate that low ability Ss had the largest number of positive interactions with their teachers, followed by the average ability groups, and then the high ability group for Grades 1-4. No corresponding F ratio was found in the Grades 3-6 analysis. While this two-factor interaction effect is more suggestive than the main effects reported previously, it is not amenable to specific interpretations. At best, it suggests that students across four grade levels have more positive interactions with their teachers if they are singled out for special treatment. This interpretation must be tempered by the fact that grade level differences were not accounted for; therefore, a con-

clusive statement cannot be made.

Since no significant achievement gains or ability gains were found, Hypotheses 4, regarding differences in instructional methods for experimental Ss when compared to control Ss, had no criterion measures on which to be evaluated. However, careful documentation of within classroom instructional methods and ability grouping revealed no significant differences between the teaching of experimental Ss and the teaching of control Ss. While differences between ability groups was noted (eg., more independent study for "brighter" students,), there was no indication that matched pairs within a given ability lever were treated differently. High ability control subjects participated in as many independent projects as did their experimental counterparts.

It should be stressed that these findings do not rule out the possibility that achievement differences reported by other investigators may have been the result of differences in instructional methods rather than experimental manipulation of teacher expectancies. The failure to find achievement differences in the present study makes it impossible to rule out the instructional methods hypothesis.

In sum, data analysis revealed that only 12 of 92 F tests computed reached acceptable levels of significance. Three, regarding over-all achievement gains, were relatively uninteresting. Of the remaining 9, 8 were generally uninterpretable main effects involving observed pupil-teacher interactions. One ability level-treatment interaction effect was reported. This effect was not amenable to concise interpretation. Thus, all experimental hypotheses were rejected, or, in the case of Hypothesis 4, untestable.

Summary and Conclusions:

The failure to find any significant effects of experimental manipulation of teacher expectations leads this investigator to concur with others in regarding the Rosenthal and Jacobson findings spurious. In a relatively well executed randomized block design, the failure to find interaction effects supporting the specific hypotheses for any of eight variables leads to the conclusion that teachers' expectations, and subsequent behavior, which may affect achievement, cannot be altered by the simple process of showing teachers cognitively inconsistent test scores. That is, while teachers' expectations may well affect the social and achievement behaviors of students, the etiological factors are varied and complex. One inconsistent test score cannot dispel other social and academic factors which impinge upon the teacher and serve as the basis for expectations.

Although some investigators have reported gains in achievement test scores as a result of expectancy inducement, the majority of negative findings suggest that significant positive findigns are either situation-specific or chance findings.

Since no significant differences in instructional methods for groups within classrooms was observed, except presenting the same material at a slower pace, hypotheses concerning modes of instruction cannot be evaluated. Furthermore, since grouping and regrouping for the control Ss and the experimental Ss was at a minimum in all classrooms, no basis for the assessment of the group placement hypothesis is available. Since no significant results in criterion measures were found, the effects of group placement and instructional methods could not be determined.

It could be argued, however, that the lack of regrouping was one of the chief reasons that there were no increases in the achievement posttest. That is, if group placement is a major cause in making the teacher's expectance come true, the failure of the teacher to regroup experimental subjects could have contributed to the lack of significant differences in achievement measures.

Based on the findings reported in this study, the other findings reported in post Rosenthal and Jacobson studies, and an analysis of theoretical positions regarding the self-fulfilling prophecy presented elsewhere (Wilkins, 1974), it is the opinion of this writer that alternative modes of investigation be pursued. It seems clear that a teacher's expectations are founded on a variety of complex and interrelated factors; and it is naive to assume that showing a teacher one or two scores on a standardized achievement test will have any effect on his original expectations. It is recommended that observational studies, following Rist (1970) and historical studies, after Seaver (1971) be undertaken in an effort to identify the various social, psychological, and academic factors utilized by teachers when they set expectations. While this writer still holds the original claim that teacher expectations affect pupil behavior and achievement, attempts to experimentally manipulate or alter those expectations are premature until the basis for and the operations of teachers' expectancies is ascertained. Candidly, it seems unlikely that we can experimentally manipulate teachers' expectations if we do not know what factors or antecedents to manipulate.

Another avenue of investigation which should be pursued was suggested by Finn (1972). Quite correctly, Finn notes that the teacher is only one

(perhaps minor) source of expectancy for the student. Parental expectations, peer expectations, and self-expectations are also impinging on the pupil's behavior and achievement. A wide variety of empirical questions are readily obvious. Is one source more potent than the others? What is the effect of incongruent expectancies from different sources? Does a change in the expectations of one source influence the child's perceptions of expectancy from other sources? Indeed, it is possible that the predominance of other sources of expectancy makes the teacher's expectations rather inconsequential.

In conclusion, it seems clear that the expectations of others play an influential role in the behavior of the individual, both in school and in the larger social arena. Furthermore, the baisis for the expectations and the individuals' perception of the various expectations are complex and generally virgin areas for investigation. Finally, it seems fair to conclude that investigations following the Rosenthal and Jacobson model have demonstrated the model to be naive, and founded on inadequate evidence. It is hoped that investigations following the ideas outlined above will be undertaken, and that these investigations will provide an understanding of a potential source of significant educational change and improvement.

[1] The first grade subjects found the Metropolitan Achievement Test too difficult and frustrating. Therefore, first grade Ss have scores only on the Analysis of Learning Potential.

ABILITY TEST SCORES
MEANS AND STANDARD DEVIATIONS FOR GRADES 1
THROUGH 6 BY ABILITY LEVEL

TABLE 1

	Experimental Ss				Control Ss			
	Pre		Post		Pre		Post	
	X	Sd	X	Sd	X	Sd	X	Sd
Grade 1								
Low Abil	85.50	13.95	108.25	7.46	84.00	13.91	107.75	9.20
Mid Abil	97.50	17.41	115.50	9.44	82.50	9.34	110.25	4.14
Hi Abil	98.00	16.62	108.50	7.53	97.50	17.14	115.50	9.44
Grade 2								
Low Abil	58.00	13.92	84.50	8.52	54.75	16.06	80.75	23.12
Mid Abil	69.00	14.30	106.50	22.87	65.25	17.35	90.25	15.13
Hi Abil	86.50	12.09	109.00	14.08	84.50	8.29	113.50	13.90
Grade 3								
Low Abil	80.50	14.70	96.50	21.61	83.25	13.26	103.00	22.07
Mid Abil	104.25	3.56	116.50	12.43	101.50	6.26	119.25	7.08
Hi Abil	117.00	8.68	113.50	9.81	114.25	7.59	136.75	2.58
Grade 4								
Low Abil	61.25	7.85	69.00	11.22	62.00	9.48	70.75	13.16
Mid Abil	61.75	8.87	58.50	8.07	61.00	8.51	71.25	5.35
Hi Abil	84.25	9.75	86.87	8.43	85.75	7.79	90.00	17.84
Grade 5								
Low Abil	67.75	8.43	83.25	16.52	61.00	2.54	82.00	10.00
Mid Abil	62.25	9.37	75.50	8.25	70.25	9.52	83.25	12.27
Hi Abil	72.75	6.72	95.25	6.97	75.50	12.59	95.00	12.58
Grade 6								
Low Abil	97.25	7.94	96.25	9.20	92.25	6.45	111.50	4.09
Mid Abil	103.75	8.61	111.25	8.87	102.75	8.25	114.25	11.18
Hi Abil	103.50	13.08	113.00	7.34	109.50	10.60	114.50	10.06

ACHIEVEMENT TEST SCORES
MEANS AND STANDARD DEVIATIONS FOR GRADES 2
THROUGH 6 BY ABILITY LEVEL

	Experimental Ss				Control Ss			
	Pre		Post		Pre		Post	
	X	Sd	X	Sd	X	Sd	X	Sd
Grade 2								
Low Achieve	49.50	17.68	88.75	20.64	49.25	23.98	91.25	32.51
Med Achieve	60.25	26.35	100.25	24.08	61.50	32.34	95.00	26.48
Hi Achieve	69.25	16.70	104.25	18.34	74.75	21.27	110.75	10.98
Grade 3								
Low Achieve	22.50	8.04	41.75	13.16	30.00	2.91	50.50	12.11
Mid Achieve	41.50	4.60	60.50	4.38	44.00	8.39	47.00	17.83
Hi Achieve	45.25	4.14	68.50	7.69	47.75	5.50	65.25	2.86
Grade 4								
Low Achieve	44.50	7.82	58.25	9.36	44.75	10.23	59.25	15.57
Mid Achieve	62.75	6.75	67.75	6.01	62.75	7.79	74.50	10.30
Hi Achieve	75.25	9.17	83.50	5.67	81.75	7.79	85.50	6.53
Grade 5								
Low Achieve	45.75	3.34	62.00	12.84	44.00	5.09	59.75	8.67
Mid Achieve	51.75	2.86	58.50	6.68	53.75	2.77	64.75	8.25
Hi Achieve	66.00	7.81	78.25	6.90	66.50	7.79	78.25	5.26
Grade 6								
Low Achieve	61.50	6.10	65.00	4.63	61.50	5.67	77.25	5.62
Mid Achieve	67.00	6.16	78.50	5.02	64.50	6.87	77.50	7.69
Hi Achieve	73.00	5.74	80.00	4.74	74.25	7.94	80.75	6.37

TABLE 2

TABLE 3
OBSERVED PUPIL-TEACHER INTERACTION
BY GRADE LEVEL AND ABILITY LEVEL

GRADE 1

	EXPERIMENTAL				CONTROL			
	Positive	Neutral	Negative	Total	Positive	Neutral	Negative	Total
Mean High Ab.	107.00	635.00	65.50	807.50	87.00	336.25	31.50	454.75
Sd High Ab.	60.09	449.58	22.16	479.54	80.90	237.24	29.26	305.57
Mean Mid. Ab.	156.75	540.25	87.25	784.25	125.50	485.50	69.00	680.00
Sd Mid. Ab.	116.70	174.81	57.62	303.21	64.43	116.18	54.94	155.60
Mean Low Ab.	117.50	793.00	118.25	974.75	1.50	135.00	5.00	141.50
Sd Low Ab.	87.90	192.23	99.38	236.06	2.59	110.79	8.66	118.26

TABLE 4
OBSERVED PUPIL-TEACHER INTERACTION
BY GRADE LEVEL AND ABILITY LEVEL

GRADE 2

	EXPERIMENTAL				CONTROL			
	Positive	Neutral	Negative	Total	Positive	Neutral	Negative	Total
Mean High Ab.	84.75	480.00	39.25	604.00	56.25	343.00	30.75	430.00
Sd High Ab.	71.74	249.93	26.65	345.28	56.83	231.91	12.89	294.51
Mean Mid. Ab.	218.00	788.50	81.75	4353.00	82.75	300.00	40.75	423.50
Sd Mid. Ab.	202.53	217.11	74.59	449.82	86.84	53.24	27.77	133.64
Mean Low Ab.	327.75	719.25	39.75	1086.75	194.00	551.75	29.00	775.25
Sd Low Ab.	248.36	346.56	34.73	519.61	198.52	276.03	13.35	421.90

TABLE 5
OBSERVED PUPIL-TEACHER INTERACTION
BY GRADE LEVEL AND ABILITY LEVEL

GRADE 3

	EXPERIMENTAL				CONTROL			
	Positive	Neutral	Negative	Total	Positive	Neutral	Negative	Total
Mean High Ab.	55.50	279.00	17.00	351.50	37.75	220.75	47.50	306.00
Sd High Ab.	56.98	103.88	14.45	100.45	41.46	96.69	74.81	173.75
Mean Mid. Ab.	68.25	271.50	13.00	352.75	43.50	341.00	49.25	433.75
Sd Mid. Ab.	73.64	73.32	10.65	113.64	43.66	50.08	56.84	33.19
Mean Low Ab.	182.00	595.50	42.50	820.00	127.50	222.36	62.25	509.50
Sd Low Ab.	226.04	195.28	50.60	380.31	184.91	222.36	89.78	408.75

TABLE 6
OBSERVED PUPIL-TEACHER INTERACTION
BY GRADE LEVEL AND ABILITY LEVEL

GRADE 4

	EXPERIMENTAL				CONTROL			
	Positive	Neutral	Negative	Total	Positive	Neutral	Negative	Total
Mean High Ab.	137.00	301.75	20.75	459.50	106.50	414.25	11.25	586.00
Sd High Ab.	124.65	126.82	17.42	260.45	89.38	69.04	11.75	63.83
Mean Mid. Ab.	121.00	509.25	49.00	679.25	77.00	255.50	28.00	372.50
Sd Mid. Ab.	44.37	284.26	37.22	288.33	44.25	75.88	12.18	90.86
Mean Low Ab.	157.25	468.00	21.00	646.25	40.00	144.50	11.50	196.00
Sd Low Ab.	91.85	303.11	14.16	330.81	23.09	73.12	5.59	52.52

TABLE 7
OBSERVED PUPIL-TEACHER INTERACTION
BY GRADE LEVEL AND ABILITY LEVEL

GRADE 5

	EXPERIMENTAL				CONTROL			
	Positive	Neutral	Negative	Total	Positive	Neutral	Negative	Total
Mean High Ab.	81.25	366.25	19.25	466.75	88.25	366.50	32.50	487.25
Sd High Ab.	65.09	80.45	17.99	155.55	63.40	154.58	26.58	167.03
Mean Mid. Ab.	110.25	452.50	22.00	584.75	62.25	313.25	31.25	406.75
Sd Mid. Ab.	123.05	206.45	17.02	337.54	48.72	62.80	23.25	103.77
Mean Low Ab.	79.00	354.75	14.50	448.25	80.00	397.00	21.25	491.50
Sd Low Ab.	115.41	188.48	18.83	314.62	122.93	170.50	21.39	309.99

TABLE 8
OBSERVED PUPIL-TEACHER INTERACTION
BY GRADE LEVEL AND ABILITY LEVEL

GRADE 6

	EXPERIMENTAL				CONTROL			
	Positive	Neutral	Negative	Total	Positive	Neutral	Negative	Total
Mean High Ab.	157.50	373.25	38.75	596.50	106.50	347.75	24.50	478.75
Sd High Ab.	98.88	137.93	62.58	234.43	155.23	225.25	25.70	368.25
Mean Mid. Ab.	142.75	351.00	32.50	526.25	137.50	342.50	47.75	527.75
Sd Mid. Ab.	120.95	98.93	44.02	240.19	114.75	226.70	56.39	379.55
Mean Low Ab.	143.25	379.50	17.50	540.25	133.25	366.25	14.00	513.50
Sd Low Ab.	100.51	191.38	24.87	306.64	143.02	241.90	11.97	388.46

TABLE 9
SOCIOMETRIC SCORES
BY GRADE LEVEL AND ABILITY LEVEL

	Experimental Ss				Control Ss			
	Pre		Post		Pre		Post	
	X	Sd	X	Sd	X	Sd	X	Sd
Grade 1								
Hi Abil	6.25	2.58	9.75	6.60	14.00	13.98	23.50	19.85
Mid Abil	2.50	2.59	4.25	2.04	14.00	11.74	11.50	8.78
Low Abil	5.25	3.36	5.25	3.03	10.75	6.75	10.50	7.36
Grade 2								
Hi Abil	5.75	6.49	3.75	4.26	12.50	5.40	8.75	6.94
Mid Abil	7.25	3.11	3.25	2.16	21.50	14.94	10.00	8.47
Low Abil	5.75	2.04	5.25	3.96	7.25	3.41	6.00	5.24
Grade 3								
Hi Abil	9.00	8.06	7.50	6.42	12.00	8.60	12.75	15.31
Mid Abil	6.25	3.26	2.00	2.12	12.75	10.63	12.00	13.91
Low Abil	13.75	17.76	4.75	2.86	4.75	3.56	1.00	1.00
Grade 4								
Hi Abil	13.25	7.66	16.50	20.22	33.50	41.40	42.25	48.58
Mid Abil	3.50	2.29	4.25	4.38	3.25	3.49	2.75	2.77
Low Abil	3.75	2.86	4.50	2.50	3.75	1.63	2.25	1.29
Grade 5								
Hi Abil	7.50	5.02	6.25	2.58	8.75	5.06	9.75	9.33
Mid Abil	2.75	1.63	3.00	1.87	3.00	4.06	4.00	4.00
Low Abil	6.75	2.16	1.75	1.29	5.25	3.76	2.25	1.78
Grade 6								
Hi Abil	6.00	5.74	6.75	6.33	21.50	24.72	12.50	16.00
Mid Abil	17.25	29.30	12.00	20.21	6.00	.70	5.25	1.08
Low Abil	3.00	1.87	6.75	6.97	1.75	2.04	3.75	2.48

REFERENCES

Anderson, D.F. and Rosenthal, R. "Some effects of interpersonal expectancy and social interaction on institutionalized retarded children." Proceeding of the 76th Annual Convention of the American Psychological Association, 3 (1968): 479-480.

Barber, T.X. and Silver, M.J. "Fact, fiction, and the experimenter bias effect." Psychological Bulletin Monographs, 70 (6, Part 2), (1968): 1-29.

Beck, E.R. "Pupil's perceptions of teacher merit: A factor analysis of five postulated dimensions." Dissertation Abstracts, 1964.

Brophy, J. and Good, T.L. "Teachers' communication of differential expectations for children's classroom performance: Some behavioral data." Journal of Educational Psychology, 61 (1970): 356-374.

Claiborn, W.L. "Expectancy effects in the classroom: A failure to replicate." Journal of Educational Psychology, 60 (1969): 377-383.

Conn, L., Edwards, C., Rosenthal, R., and Crowe, D. "Emotion perception and response to teacher expectancy in elementary school children." Psychological Reports, 22 (1968): 27-34.

Elashoff, J. and Snow, R. Pygmalion Reconsidered. Worthington, Ohio: Chas. Hones Publishing Co., 1971.

Evans, J. and Rosenthal, R. "Interpersonal self-fulfilling prophecies: Further extrapolation from the laboratory to the classroom." Proceedings of the 77th Annual Convention of the American Psychological Association, 4 (Part 1) (1969): 371-372.

Finn, J. "Expectations and the educational environment." Review of Educational Research, 42 (1972): 387-410.

Flanagan, N. and Havumaki, S. "The effects of teacher-pupil contacts involving praise on the sociometric choices of students." Journal of Educational Psychology, 51 (1960): 65-68.

Fleming, E. and Anttonen, R. "Teacher expectance, or my fair lady." American Educational Research Journal, 8 (1971): 241-252.

Fox, R. Lippitt, R., and Schmuck, R. Pupil-Teacher Adjustment and Mutual Adaptation in Creating Classroom Learning Environments. Final Report, Cooperative Research Project, No. 1167, Univ. of Michigan, 1964.

Goldsmith, J.S. and Fry, E. "The test of a high expectancy prediction on reading achievement and IQ or students in grade 10 (or, Pygmalion in puberty)." Submitted to the American Educational Research Journal, 1970.

Good, T.L. "Which pupils do teachers call on?" The Elementary School Journal, 70 (1970): 190-198.

Jose, J. and Cody, M. "Teacher-pupil interactions as it related to attempted changes in teacher expectancy of academic ability and achievement." American Educational Research Journal, 8 (1971): 39-49.

Kester, S. "The communication of teacher expectations and their effects on the achievement and attitude of secondary school students." American Educational Research Journal, 66 (1972): 51-55.

Lippitt, L. and Gold, M. "Classroom social structure as a mental health problem." Journal of Social Issues, 15 (1959): 40-58.

Meichenbaum, D., Bowers, K., and Ross. "A behavioral analysis of teacher expectancy effect." Journal of Personality and Social Psychology, 13 (1969): 306-316.

Palardy, J. "What teachers believe, what children achieve." Elementary School Journal, 69 (1969): 370-374.

Pellegrini, R. and Hicks, R. "Prophecy effects and tutorial instructions for the disadvantaged child." American Educational Research Journal. 9 (1972): 413-420.

Rist, R. "Student social class and teacher expectations; The self-fulfilling prophecy in ghetto education." Harvard Educational Review, 40 (1970): 411-451.

Rosenthal, R. and Jacobson, J. Pygmalion in the Classroom. New York: Holt, Rinehart and Winston, Inc., 1968.

Rothbart, M., Delfen, and Barrett, F. "Effects of teacher's expectancy on student-teacher interaction." Journal of Educational Psychology, 62 (1972): 49-54.

Rubovits, P. and Haehr, M. "Pygmalion analyzed toward an exploration of the Rosenthal-Jacobson findings." Paper presented at Annual Meeting, American Educational Research Association, Minneapolis, 1970.

Schrank, W.R. "The labeling effect of ability grouping." The Journal of Educational Research, 62 (1968): 51-52.

Schrank, W.R. "Further study of the labeling effects of ability grouping." The Journal of Educational Research, 63 (1970): 358-360.

Seaver, W.B. "Effects of naturally induced teacher expectancies on the academic performance of pupils in primary grades." Doctoral Dissertation University of Illinois, 1971.

Thorndike, R.L. "Review of Pygmalion in the Classroom." American Educational Research Journal, 5 (1968): 708-711.

Wilkins, W.E. "The concept of the self-fulfilling prophecy." Paper accepted for publication, ~~Society~~ Sociology of Education, 1974a.

Wilkins, W. E. "Some explanations for the failure to replicate the Rosenthal and Jacobson findings." Paper under review, 1974b.