CUEING—OPERANT CONDITIONING:
MEDIATORS OF EXPERIMENTER EXPECTANCY?

by

MICHAEL CARTER MOFFAT
B.Sc., The University of British Columbia, 1964

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS

in the Department
of
Psychology

We accept this thesis as conforming to the
required standard

THE UNIVERSITY OF BRITISH COLUMBIA
December, 1967
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Psychology
The University of British Columbia
Vancouver 8, Canada

Date December 19, 1967
The psychological experiment is frequently regarded as a situation which allows for complete control over the inputs to the experimental Ss. Closer examination reveals it to be vulnerable to unwanted and unprogrammed experimenter influence, mediated through essentially two modes of communication—visual-kinesic and auditory-paralinguistic.

This study examines the effect of different types of S to E feedback of information, mediated by verbal and nonverbal channels of communication, upon the E outcome-bias phenomenon. Experimenters who were given an expectancy for certain responses from their Ss, were placed in an experimental situation that permitted or restricted verbal communication, and included correct, reversed, or no feedback of subjects' responses, on a photo-rating task.

Ten Es each ran 12 Ss on a photo-rating task in a study purporting to be a research project developing a test of empathy. The Ss were required to examine 20 standardized neutral photographs of faces and to rate each one on the degree of success or failure that the person pictured had been experiencing. The Es had been led to expect a predominance of success responses from their Ss.

The principal hypothesis, that E expectancy effects are independent of the type of S to E feedback, was supported: the photo-ratings by Ss showed a significant increase in magnitude as a function of the number of photos rated, irrespective of the type of S to E feedback permitted. The more photos rated, the greater the magnitude of the success rating. This has
relevance for the type of process that should be considered as a mediator of 
E's expectancy. Three alternative processes are discussed in the light of 
the findings of the present study.

The second hypothesis, that greater E bias effects occur in conditions 
permitting both verbal and nonverbal cues as compared to conditions per­
mitting nonverbal cues alone, was not supported. This suggests that verbal 
cues do not make a significant contribution to E bias effects during the data 
collecting phase of the psychological experiment.

Implications of the findings of this study and suggestions for future 
research are discussed.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I The Problem and Definitions of Terms Used</td>
<td>1</td>
</tr>
<tr>
<td>The problem</td>
<td>1</td>
</tr>
<tr>
<td>Definitions of terms used</td>
<td>3</td>
</tr>
<tr>
<td>II Review of the Literature</td>
<td>4</td>
</tr>
<tr>
<td>Evidence for the occurrence of experimenter effect</td>
<td>4</td>
</tr>
<tr>
<td>Contribution of experimenter characteristics to the experimenter bias phenomenon</td>
<td>6</td>
</tr>
<tr>
<td>Experimenter expectancy as a determinant of experimental results</td>
<td>7</td>
</tr>
<tr>
<td>Communication of experimenter bias</td>
<td>9</td>
</tr>
<tr>
<td>Modality of cue communication</td>
<td>14</td>
</tr>
<tr>
<td>Factors affecting the bias phenomenon</td>
<td>17</td>
</tr>
<tr>
<td>Concluding remarks</td>
<td>19</td>
</tr>
<tr>
<td>III Method</td>
<td>27</td>
</tr>
<tr>
<td>IV Results</td>
<td>35</td>
</tr>
<tr>
<td>V Discussion</td>
<td>48</td>
</tr>
<tr>
<td>VI Summary and Conclusions</td>
<td>57</td>
</tr>
<tr>
<td>Bibliography</td>
<td>61</td>
</tr>
</tbody>
</table>

APPENDICES

A Standardization procedure for the person perception task | 65 |
B Instructions to subjects in the Verbal Correct Feedback
C Instructions to subjects in the Verbal Reversed Feedback condition .................. 67
D Instructions to subjects in the Verbal No Feedback condition .................. 68
E Instructions to subjects in the Nonverbal Correct Feedback condition .................. 69
F Instructions to subjects in the Nonverbal Reversed Feedback condition .................. 70
G Instructions to subjects in the Nonverbal No Feedback condition .................. 71
LIST OF TABLES

TABLE PAGE

I Experimental Design: 2 x 3 x 10 Factorial ................. 30
II Experimenters' Mean Photo-Rating Scores for the Six
    Experimental Conditions .................................. 35
III Mean Photo Ratings Over Blocks of Trials For Each of the
    Six Treatment Conditions ................................ 36
IV Summary of the Four Factor Repeated Measures ANOVA for
    the 20 Trial Photo-Rating Task ............................. 37
V Trend Analysis of the Grand Means for the Four Trial
    Blocks ........................................................... 39
VI Comparisons of the Grand Means for the Four Trial Blocks .... 40
VII Analysis of Variance of Obtained Photo Ratings for Group NVR .. 41
VIII Analysis of Variance of Obtained Photo Ratings for Group NVK .. 42
IX Analysis of Variance of Obtained Photo Ratings for Group NVC .. 42
X Analysis of Variance of Obtained Photo Ratings for Group VC. 43
XI Analysis of Variance of Obtained Photo Ratings for Group VR. 43
XII Analysis of Variance of Obtained Photo Ratings for Group VK. 44
XIII Summary of the Analysis of Variance on the First Ten Trials.. 45
XIV Trend Analysis of the Grand Means for the First Nine Trials.. 46
# List of Figures

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Eight Point Success-Failure Rating Scale for the Correct Feedback Condition as Seen by Both E and S</td>
<td>29</td>
</tr>
<tr>
<td>2. Eight Point Success-Failure Rating Scale for the Reversed Feedback Condition. A, as Seen by E; B, as Seen by S</td>
<td>29</td>
</tr>
<tr>
<td>3. Grand Mean Photo Ratings Obtained for the Four Trial Blocks</td>
<td>38</td>
</tr>
<tr>
<td>4. Grand Mean Photo Ratings Obtained for the Twenty Trials</td>
<td>47</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENT

The writer would like to express his gratitude to those individuals whose cooperation made it possible to accomplish this study. In particular, thanks are due to Dr. R. E. Knox and Dr. R. Wong who provided valuable constructive criticism of the design and writing of this study. Others who also contributed were: Carole Hilstrom, who provided expert typing and secretarial assistance; and Keith Waldron, with his technical advice in matters concerning apparatus design.
CHAPTER I
THE PROBLEM AND DEFINITIONS OF TERMS USED

One source of unwanted variance in psychological research that has sparked a great deal of interest in the past decade has been the experimenter himself. Although the problem of experimenter contamination perplexed scientists as early as 1911 when Pfungst documented the effect of experimenter influence in determining the behavior of Clever Hans, the horse that could "solve" all manner of problems, little has been done until recently in the way of studying the experimenter as an independent variable. A finding of these recent studies is that uncontrolled and unconscious experimenter input into the S-E relationship may affect the outcome of the psychological experiment. The experimenter who expects a particular outcome can, through covert verbal and visual cues, unintentionally influence his Ss' responses in accordance with his expectancy. It has been suggested that verbal conditioning may play a significant role in the mediation of the experimenter's bias. There are, however, very little experimental data bearing on this issue. Evidence to confirm or to refute this suggestion has been inconclusive.

The Problem

Statement of the problem. The purpose of the present study is threefold: (1) to compare the effects of permitting or eliminating auditory cues from the S-E relationship upon E bias, (2) to determine if the type of communication from Ss to E affects the Ss' responses, and (3) to investigate whether verbal or nonverbal conditioning is or is not in
operation in the communication of the E's bias.

**Importance of the study.** The laboratory is a preferred mode for the observation of behavior. The reason for this preference is due to the greater control it provides over the inputs to the experimental Ss. Some measure of control over the inputs is lost when E communicates unintentionally and differentially with his Ss. Because such control is the reason for the reliance on the experimental method, there are serious implications. Some attempts have been made to reduce E effects, they include: the employment of naive research assistants; the use of the "double blind" method in which no one having direct contact with Ss is permitted to know what the Ss' treatment condition will be, is or has been, until the experiment is over; the tape recording of Ss' instructions; and the elimination of verbal and/or visual communication between E and Ss. In spite of these and other techniques devised to reduce E's contamination of his own data, his unwanted influence often persists. The limitations of these techniques have been pointed out by Kintz, Delprato, Mettee, Parsons, and Schappe (1965), Rosenthal (1963c, 1965, 1966).

Prior to actually implementing control measures aimed at reducing E effect, three elements must be identified: (1) the point in the experimental procedure at which the contamination takes place, (2) the mode of E contamination, and (3) the specific cues used in the contamination. Many attempts to eliminate E effects have instituted control techniques without first identifying these elements. The lack of economy of such an approach is obvious. What is needed is a careful examination of these issues so that researchers may devise less omnidirectional and more effective control techniques.
Such an examination is the general purpose of this study.

Definitions of Terms Used

**Experimenter expectancy.** This term will refer to the expectancy that E has concerning the outcome of the experimental results. It may originate with E or it may be induced, in the case where one or more assistant Es are employed by a senior E.

**Experimenter outcome-orientation bias.** Experimenter bias will be interpreted as the phenomenon that is in operation when Es obtain from their Ss, human or animal, the data that they want and/or expect to obtain which are asymmetrically distributed about the "correct" or "true" value, cf. Rosenthal, 1966.

**Covert communication.** Covert communication refers to either verbal or nonverbal communication of which neither the communicator nor the receiver of the communication are aware.

**Correct feedback.** This term refers to the transmission of Ss' true responses to E.

**Reversed feedback.** This term refers to the transmission of Ss' true responses to E in such a way that E unwittingly interprets the response as meaning the opposite to that intended by Ss.

**No feedback.** No feedback refers to the situation in which Ss' responses are not revealed to E.

**Cue effect.** A cue effect refers to data biasing by E through his use of distinctive but covert cues which communicate to Ss the type of response that is desired. The cues may or may not be rewarding to Ss.
CHAPTER I
REVIEW OF THE LITERATURE

Much of the literature in regard to E effect contamination has focused on the source of E bias, its mediation, and the variables affecting E bias; but only the work done on the problems very closely related to the one at hand will be summarized here.

Evidence for the Occurrence of Experimenter Effect

Any serious study of E effect must not overlook the possibility that the effects under investigation were intentional. Rosenthal (1966) cites several cases where experimental results have been fabricated. The occurrence of such cases, it must be assumed, is rare, however. The problem to be considered in the present discussion is an equally important and more frequent one, the problem of unintentional E effects.

Rosenthal (1966) makes a distinction between observer errors or effects and observer "bias". He regards observer errors as being randomly distributed around a "true" value. Biased observations, on the other hand, are regarded as reflecting a trend. This trend may be a function of some characteristic of the observer, or the observation situation, or both. Since biased observations do not occur randomly, they are to some extent predictable. Current research on E bias is concerned with identifying relationships between observer and situation characteristics and biased observations. The consequence of this endeavour, it is hoped, will lead to accurate and efficient methods of predicting biased observations.

Recent experiments have shown that different Es obtain different
results on similar experimental tasks using comparable Ss (Harris, Piccalino, Rodback and Sommer, 1964; Lord, 1950; and Postman and Jarrett, 1952).

Using a verbal conditioning paradigm, Kanfer (1958) reinforced verb responses with a flashing light under three reinforcement schedules. Two similar Es were employed. In reinforcing Ss, E was required to distinguish between verbs and nonverbs. Even with this simple task, a significant interaction was found between E and method of reinforcement. That is, Es differed in their tendency to reinforce Ss within the treatment conditions even though both Es received the same instructions on how to reinforce Ss.

A recent study by Severin and Rigby (1963) investigated different patterns of digit groupings on the recall of seven digit numbers. In analyzing the variance of correctly recalled numbers an E effect was found significant at the .01 level. On further analysis it was found that this effect was due largely to one of the four Es involved. A repetition of the analysis without this particular E’s data yielded no significant E difference.

Nash, Phelan, Demas, and Bittner (1966) studied the effects of manifest and induced anxiety and E variability on reaction time. Thirty-six female Ss were placed in low, medium or high anxiety groups based on their scores on the Taylor Manifest Anxiety Scale. The Ss were then randomly assigned to either stress or no stress treatments and to one of two Es. Simple reaction times were recorded for all Ss. Significant performance differences were found when Es alternately administered the experimental treatment.
These experiments are examples of studies which have shown non-differentiated E influence, that is, an influence that cannot be ascribed either to unintended differences in E's behavior or to differences in his appearance. The number of studies reporting such influence is few as are attempts at discerning the possible reasons for the differing data. Recent attention to E effects has, however, prompted researchers in the physical, biological and behavioral sciences, to be more cognizant of the pervasiveness of the problem. As a consequence of this new awareness, concerted efforts are being made to identify the factors contributing to E effects.

Contribution of Experimenter Characteristics to the Experimenter Bias Phenomenon

Reviews of E attributes that have been found to be partial determinants of the responses given to E by his Ss in a variety of data-collecting situations have been prepared by Kintz, et al. (1965) and Rosenthal (1963a, 1965, 1966). Among the attributes that have been studied are E's race, religion, status, likeability, warmth, sex, personality, anxiety level, social desirability, experience, and acquaintance with Ss. Rosenthal (1963c) suggests that a more superordinate set of constructs must yet be empirically developed to account for the findings relating to these variables. Pending such a development, he concludes that Es will obtain differential data as a function of how E is regarded by his Ss in the attributes of (1) likeability, (2) prestige, (3) professional skill, (4) trust, and (5) sex.
Experimenter Expectancy as a Determinant of Experimental Results

Perhaps the most disturbing phenomenon attributable to E's presence in the experimental situation is his ability to influence his Ss to yield data consistent with his hypothesis. Upon investigation, it has been demonstrated that the expectancy with which an E enters the laboratory is a significant determinant of his results (Rosenthal and Fode, 1960, 1963a, 1963b; Rosenthal, Fode and Vikan-Kline, 1960; Rosenthal, Persinger, Vikan-Kline and Fode, 1963a). The basic paradigm for such studies has been to divide a sample of Es into two groups and to create in each, an expectation for the data they will obtain. The expectancy given to one group of Es is opposite to the expectancy induced in the other. Both groups of Es then run comparable Ss on a similar experimental task.

In an experiment purported to concern learning in rats, Rosenthal and Fode (1963a) randomly assigned rats to two groups of Es. One group of six Es was instructed that its group of rats was "maze-bright" and a second group of six Es was instructed that its group of rats was "maze-dull". In a simple T maze, the "maze-bright" rats performed significantly better than the "maze-dull" rats. It was suggested that differences in handling of the animals was the basis for the differences of obtained experimental results. Post experimental questionnaires indicated that Es expecting good performance from their Ss had rather positive attitudes toward them. This was characteristic of Es with "maze-bright" rats. It is believed that these positive attitudes may have influenced the quality and quantity of handling behavior of the "maze-bright" rats. This interpretation was supported to some extent by the questionnaire responses of Es and by
Informal observations made during the experiment. The Es running "bright" rats saw themselves as handling their rats more gently. Support for the hypothesis that increased handling can improve performance is given by Rosenthal and Lawson (1964), Bernstein (1952) and Bernstein (1957).

In the Rosenthal and Lawson (1964) study two groups of Es, one with supposedly bright rats, the other with supposedly dull rats, carried out seven experiments. These included tasks such as magazine training, operant acquisition, stimulus discrimination, and chaining of responses. The Es who believed their rats to be bred for brightness obtained results that significantly differed on seven out of the eight comparisons in the expected direction from results obtained by Es who believed they had dull rats. It was suggested that these differences may be due to differential reinforcing effects of the Es' handling behavior. Descriptions of Es' handling behavior before and after the experiments revealed that the Es expecting better performance handled their animals about 33% more after each experiment, whereas Es expecting poorer performance handled their animals about 44% less after each experiment.

The effect of E expectancy has also been examined in studies utilizing human Ss. Robert Rosenthal has conducted a number of person perception experiments in which Es were to obtain ratings of photographs on a success-failure dimension. The photos were of faces cut from a weekly news magazine and selected so that under standard conditions most Ss rated them neutral on a 20 point success-failure scale. In two different experiments (Rosenthal and Fode, 1961; and Fode, 1965) Es in one group were told that they would probably obtain mean ratings of +5 from Ss while
Es in another group were told that they would probably obtain mean ratings of -5 from their Ss. In these studies, the lowest mean ratings obtained by any E expecting high ratings was higher than the highest mean ratings obtained by any E expecting low ratings from his Ss. This finding not only illustrates the existence of E bias but the possible magnitude of the problem in experimentation.

Communication of Experimenter Bias

Two phases of the experimental situation have been studied in an attempt to isolate when E's expectancy is communicated to Ss in the S-E interaction. These phases are (1) the predata collecting phase during which time E greets and instructs Ss, and (2) the data collecting phase when E administers the experiment and collects the data.

Rosenthal (1966) provides evidence to support his contention that Es may communicate their expectancies to Ss in both phases of the experiment. Two films were made of E-S interactions during a photo-rating task. Three E-S interactions were contained in one film and five were contained in the other. In both films, expectancies between +10 (extreme success) and -10 (extreme failure) were randomly assigned to Es. The Es' task was to try to influence their Ss to rate the standard photos used in the person perception studies in the desired way (in accordance with Es expectancy) but without being too obvious. The first film was viewed by 52 observers, the second film was viewed by 11 observers. The observers were primarily university faculty members and graduate students. Their task was to try to identify Es' expectancy and to state the reason for
their judgement. The median correlations between the observers' identification of Es' expectancies and the actual expectancies for the first and second film were +.88 \( (p < .00001) \) and +.72 \( (p < .001) \) respectively. The evidence is unequivocal as to the accuracy of observers in identifying Es' expectancies when these are being deliberately communicated.

The hypotheses given by observers fell along two dimensions: temporality and sense modality. "About half the hypotheses emphasized Es' reactions to Ss' responses. For these observers, expectancy communication occurred only after Ss began responding and followed a differential reinforcing paradigm. For the other observers, expectancy communication occurred before the S even made his first response" (Rosenthal, 1966, p. 284). These hypotheses define the temporality dimension. The sense modality dimension was defined primarily in terms of hypotheses about visual and auditory channels of communication. Observers emphasized the importance of speed and manner of instruction reading, number of errors made in reading instructions, and pauses used to accentuate the desired region on the rating scale as auditory cues used by Es to communicate their expectancies. The visual cues that were hypothesized as being communicators of expectancies were more numerous, and included behaviors such as smiling, head shaking, head nodding, raising eyebrows, "looking happier", pencil tapping, and holding up the photo longer. Although these cues were found to be associated with the magnitude of expectancy effects, there still remains the questions of how Es with opposite expectancies differentially influence the responses of their Ss, a matter that will be dealt with later in this chapter.
It must be remembered that the Rosenthal (1966) study was a contrived effort, the purpose of which was to simulate the real person perception study and to determine if observers could identify the mode and time of E expectancy communication. The Es were purposely trying to bias their Ss. One would expect that cues used by Es in this situation would reflect on a molar level those cues used by Es in a real person perception study. If this is the case, subsequent investigations must attempt to refine techniques of observation and focus on subtle expressions of the molar cues that have already been identified as communicators of E expectancy. At the present time, further film analyses are being undertaken by Rosenthal in the hope that more specific modes of communication can be revealed.

Additional data exist bearing on the more molar aspects of E bias communication. These will be described in terms of the temporal phases that have been delineated above.

The predata collecting phase. Evidence of the communication of E's expectancy during the brief predata collecting phase of the S-E interaction was found in a study by Rosenthal, Fode, Vikan-Kline and Persinger (1964). In analyzing the data from three experiments on E bias they found that the bias phenomenon was in evidence on the Ss' very first response. This implies that some bias is communicated to Ss prior to the actual administration of the task.

A study by Weick (1963) lends support to the hypothesis of E bias communication during the predata collecting phase of the experiment. He employed two Es, each of whom administered the photo rating task to five introductory psychology students. The E expecting high ratings obtained
higher mean ratings than did E expecting low ratings. Careful examination of the data revealed a significant E effect on the first photo (p < .02). The fact that E expectancy effects were manifested on the ratings of the very first photo suggests that the brief period during which E greets, seats, and instructs his Ss provides time enough for E to communicate his expectancies to Ss.

Friedman, Kurland and Rosenthal (1965), investigating the effects of E behavior on experimental results, found that those Es who obtained data more in accordance with their expectancy differed in both their vocal and nonvocal behavior during the instruction reading period from those who did not obtain data in accordance with their expectancy. The Es who biased the responses of their Ss were found to read their instructions more rapidly and more accurately and to exchange fewer mutual glances with their Ss during the instruction reading phase of the experiment. In summary, it would seem that the E with a more professional manner is more likely to exert his influence on Ss than is the E who lacks such a professional manner.

A subsequent study in a person perception context, recorded S-E interactions on sound motion pictures (Rosenthal, Friedman and Kurland, 1966). Five male and five female Es were led to expect about half of their Ss to give ratings on a photo-rating task averaging +5 (success) and to expect ratings of -5 (failure) from the remaining Ss. Regardless of which of the two contradictory expectancies they held, male Es obtained significantly more hypothesis confirming data if they read instructions more rapidly and less accurately, glanced less at their Ss, exchanged fewer
mutual glances with them, and showed less general body movement. The instruction reading behavior of female Es was less predictive of their biasing effect upon their Ss' responses.

The data collecting phase. Studies examining the effects of E bias during the administration of the data collecting phase of the experiment have focused primarily on the verbal conditioning phenomenon. Greenspoon (1955) has demonstrated the effect of verbal reinforcement by reinforcing plural nouns with "mmm-hmm". Verbal reinforcement resulted in an increase in the frequency of emission of the plural nouns. By the same principle, it would not seem unreasonable to suggest that an E who expects to obtain high ratings on photographs may subtly reinforce that type of response thus increasing the probability that the response will be repeated. Conversely, if E expects low ratings, he may subtly reinforce those ratings that are low. In the study by Rosenthal, Fode, Vikan-Kline and Persinger (1964) cited above, data from three experiments on E bias were analyzed, to determine the extent that E expectancies were mediated by verbal conditioning. The results showed that verbal conditioning was not necessary in the mediation of the E bias phenomenon. In fact, as earlier reported, the bias phenomenon was already in evidence on the Ss' first response. In this study, the magnitude of E bias was defined as the difference in mean ratings obtained from two groups of Es with opposite expectancies. It may be argued that since a significant bias effect was found on the ratings of the first photograph, the ratings were already at their extreme points on the scale. This would explain why no subsequent increase in bias effect was found. If verbal conditioning was mediating the E bias, one would expect an
increase in the effect from the first rating to some asymptotic level as a function of the number of trials employed. Since this asymptotic level may already have been reached because of events in the predata collecting phase of the experiment, it is expected that verbal conditioning would not have been in operation during the data collecting phase. According to Rosenthal, "Verbal conditioning is neither a necessary condition nor a necessary augmenting factor in the operation of the E outcome-bias phenomenon" (Rosenthal et al., 1964, p. 74). This conclusion does not exclude the possibility that verbal conditioning can mediate E outcome-bias. A true test of the effects of verbal conditioning in the mediation of E bias in the person perception type of study should eliminate biasing effects prior to the operation of the independent variable. That is, there should be no evidence of E bias on ratings of the first photograph. This may be achieved by limiting the contact between E and Ss during the predata collecting phase of the experiment.

A study by Fode, Rosenthal, Vikan-Kline and Persinger (1961) adds weight to the hypothesis that verbal conditioning may be an important factor in the biasing of experimental results. They demonstrated that operant conditioning could drive the ratings of the photos up or down according to the will of E. Thus Rosenthal (1963c) concludes that, "While verbal conditioning is neither a necessary nor a necessarily frequent antecedent of biasing, it nevertheless could be" (p. 275).

Modality of Cue Communication

The question now arises as to how E unintentionally and different-
lally communicates with his Ss. What are the cues and how are they transmitted? In an attempt to answer this question, Fode (1960) placed a group of Es behind screens to eliminate visual cues, while another group of Es was required to remain silent throughout the experiment to eliminate verbal cues. Restriction of visual cues accounted for about 80% of the observed bias.

In 1963b Rosenthal and Fode conducted two experiments designed to investigate the transmission of cues from E to Ss. They used the person perception experiment employing their photo-rating task. In both experiments some Es were given an expectancy that their Ss would yield a mean photo rating of +5 (success) and the other Es were given the -5 (failure) expectancy on a success-failure rating scale. In the first experiment Es were not permitted to say anything to their Ss other than what was on the instruction sheet. Since there was no verbal communication during the photo presentations the resulting bias could only have been communicated through manner, gesture, or facial expression. In the second experiment, a different method of photo presentation was employed. The photos were mounted on cardboard and numbered so that Ss could rate them without E handling the photos. Two groups of Es served as replicates for the high (+5) and low (-5) bias groups of Es of the first experiment. Both of these groups were employed as control groups in the second experiment. A third group of Es, non-visual (+5) bias, greeted their Ss, as did all Es, but then immediately sat down behind a screen across a table from Ss. The E read instructions to Ss and recorded Ss' ratings but was never visible to Ss. The Es had been biased to expect high +5 ratings. The
fourth group of Es, non-verbal (+5) bias, after greeting Ss, handed them a sheet of instructions and remained completely silent until the end of the session. These Es were in full view of Ss. It was found that the elimination of visual cues from E to S did significantly reduce the effect of Es' bias. Elimination of verbal cues had an even more profound effect in reducing E bias. It is important to note here that the elimination of verbal cues included the elimination of E's instruction reading, whereas in the condition eliminating visual cues, instruction reading behavior was not eliminated. It has already been demonstrated that instruction reading behavior may be sufficient in itself for mediating E bias (Friedman, Kurland and Rosenthal, 1965; Rosenthal, Fode, Vikan-Kline and Persinger, 1964; Rosenthal, 1966; Weick, 1963). Thus, the greater reduction in E bias effect in the condition which eliminated verbal cues may be partially attributable to the absence of the effects of instruction reading.

The two preceding studies (Fode, 1960; and Rosenthal and Fode, 1963b) demonstrate that both verbal and visual cues play an important part in the phenomenon of E bias.

Rosenthal (1965) reports studies designed to identify the specific cues that mediate E bias. In these studies, person perception experiments utilizing the photo rating task were recorded on film. The Es who expected their Ss to be "success-perceivers" communicated their expectancy so that these Ss became "success-perceivers". Similarly, Es expecting their Ss to be "failure-perceivers" obtained a predominance of failure responses. No specific expectancy mediating cues could be identified in the three years that the films were analyzed. It may be recalled
that Es who manifested particular behavior characteristics (i.e., read instructions more rapidly and less accurately, glanced less at their Ss, exchanged fewer mutual glances with them, etc.) obtained significantly more hypothesis confirming responses from their Ss. The problem still remains, however, of determining how E uses these and other specific cues for differentially communicating his expectancies to Ss.

Factors Affecting the Bias Phenomenon

It is not unreasonable to suggest that the mood of E is governed in part by the success or failure of his ongoing research. If the researcher begins to obtain data contrary to his expectations, might not his subsequent results differ from the researcher who's expectancies are initially confirmed on the same task? Rosenthal, Persinger, Vikan-Kline and Fode (1963b) provide evidence that suggest that E's mood may be perceived by his Ss and this in turn may affect the experimental results.

In that study, three groups of four Es each had three groups of Ss rate the apparent success of people in photos on a scale running from -10 to +10. The Es were instructed that Ss' mean ratings would be about +5. In each of the two experimental groups were two confederates of the Principal Investigator. One pair of confederates was instructed to give "good data" (confirm E's expectancy) and the other pair was to give "bad data" (respond contrary to E's expectancy). It was hypothesized that Es who obtained "good data" would obtain "better" subsequent data while Es expecting "bad data" would obtain "worse" data in relation to the control. While neither experimental group differed significantly from a control
group, the two experimental groups differed significantly from each other as predicted.

The fact that E's expectancy was confirmed or disconfirmed in the preceding study may be irrelevant, however, to the state of E's mood and its effect on subsequently obtained data (Rosenthal, Kohn, Greenfield and Cartoa, 1965). Rosenthal (1966) offers an alternative explanation; that is, E's expectancy may change during the course of the experiment which may in turn result in changes in the data he obtains from his Ss. Thus, Es who have had their expectancies confirmed in the early stages of the experiment may expect similar data from future Ss. On the other hand, Es who received data from their first run Ss that disconfirmed their expectancies may begin to have doubts about their initial expectancies and modify them so as to be more consistent with the data they have already obtained. Differences of E's expectancies then, rather than differences of mood may account for the findings in the preceding study.

This discussion directs our attention to the idea that Ss may indirectly affect their own responses or the responses of subsequently run Ss. An extension of this idea may be used to explain how E learns to communicate his expectancy. If the type of cues used by E to communicate his expectancy to Ss is contingent upon the nature of Ss' responses (i.e., expectancy confirming or disconfirming responses) then Ss may unwittingly shape E's behavior.

Rosenthal (1965) points out that since it is rewarding to have one's expectancies confirmed, then; "Whenever the subject responds in accordance with the experimenter's expectancy, the likelihood is increased
that the experimenter will repeat any covert communicative behavior which may have preceded the subject's confirming response. Subjects, then, may quite unintentionally shape the experimenter's unintended communicative behavior" (p. 31). No systematic investigation of Rosenthal's notion has been carried out, however.

Concluding Remarks

In the foregoing discussion, evidence was presented which suggested that E's expectancy could be communicated during the brief predata collecting phase of the experiment. The cues involved in the communication were believed to be both verbal and visual in nature. Although observers have not been able to identify the specific cues used by E to communicate his expectancies, it would seem that E is covertly saying to his Ss, "Give me response X", and his Ss somehow tend to comply. In other words, E may in effect be cueing his Ss as to how they should respond.

The unique feature about the concept of a cueing process mediating E expectancy stems from the observation that Ss give biased ratings on the first trial of the person perception studies. It follows, then, that cueing is not dependent for its operation upon the feedback of responses from Ss to E. Cueing is a unilateral mediator of information, mediating informational cues about E's expectancy to Ss. The unilateral communication feature of cueing is not shared by the mediating process of verbal operant conditioning. Operant conditioning, of course, depends upon reciprocal communication between E and Ss. Both cueing and verbal conditioning may mediate E expectancy during the data collecting phase of the experiment.
when reciprocal communication between E and S is usually permitted. The problem now arises as to which of the two processes is most important in mediating E expectancy. That is, cueing and not verbal conditioning may be the mediator of E expectancy during the data collecting phase of the experiment. The inconclusive reports as to the significance of verbal conditioning as a mediator of expectancy warrants a careful examination of the cueing process in expectancy communication during both the predata collecting and the data collecting phases of the experiment.

It is interesting to compare and contrast the mechanisms of verbal conditioning and cueing in the role of expectancy mediation. First, let us consider Rosenthal's (1965) suggestion that Ss may unintentionally shape their E's unintended communicative behavior. The Es may be rewarded by having their expectancies confirmed. As a consequence they repeat the communicative behavior which immediately preceded the Ss' hypothesis confirming response. That is, Rosenthal agrees that verbal conditioning may contribute to the bias effect, but, he regards it as a mechanism by which Ss shape their E's behavior.

In previous studies that have investigated the bias effect as a verbal conditioning phenomenon, the communication from E was assumed to be the reinforcement for Ss' responses. The alternative suggested by Rosenthal (1965) would also involve communication from E to Ss, but the communication, as he regards it, contains distinctive cues (hints or suggestions as to what to do or when to act) that affect the way Ss respond. The cues may or may not be reinforcing for Ss. What is more important, is the informational content contained in the cues.
One may expect that the total information provided by E's cueing behavior would be a function of the length of time he spends with his Ss; the greater interaction time or the longer task results in a more refined and efficient communicative net. As a consequence, one would expect an increase in bias magnitude over trials, as is found in studies of verbal reinforcement. The mechanism involved, however, would be very different.

Rosenthal's hypothesis seems to be related to the cueing hypothesis. Both involve E's covert communication of his expectancy prior to the Ss' response. The mechanism that Rosenthal describes differs in that biasing attributable to it cannot occur until after the Ss give their first response. Moreover, E's communicative behavior is largely determined by Ss' responses. The effect of cueing (cue effect), on the other hand, can occur prior to the Ss' responding and may be independent of the type of responses Ss give. The cueing process is persistent, whereas, the effect proposed by Rosenthal depends upon E's acquiring the appropriate cues as the result of a conditioning process.

It is difficult to distinguish between reinforcing effects and cueing in trying to account for changes in Ss' behavior. The effects of cueing and operant conditioning are complementary in studies of verbal conditioning. Experimenters, reinforcing personal pronouns with the words "Good" or "Ummm Hmmm", may also be communicating distinctive cues which influence Ss' behavior so that they respond in accordance with the experimental hypothesis. One purpose of the present study was to attempt to determine the relative contribution of cue effects and reinforcing effects to changes in Ss' performance. In addition, the study was concerned with the
identification and operation of variables associated with experimenter effects.

A preliminary attempt was made to distinguish between cue effects and reinforcing effects in a person perception study (Moffat, 1966). In this study, E-S interaction during the predata collecting phase was restricted in order to prevent bias effects prior to the administration of the task. The aim of this procedure was to avoid any communication of bias prior to presentation of the first photo.

In order to distinguish between the effects of covert reinforcement and cueing, a procedure was adopted by which the effects of these complementary mechanisms could be isolated. This was achieved by using different types of feedback from S to E: correct feedback, reversed feedback, and no feedback. Ten Es each ran two Ss in each of the feedback conditions. The Es were told that they would be running Ss selected because they were prone to rate the photos "success". The Es' task was to administer the 20 standardized photos used by Rosenthal (1966) and to obtain ratings of the photos of either "success" or "failure" from Ss. Both E and Ss were instructed to say nothing prior to or during the administration of the task. Thus, verbal cues were eliminated as a possible mode for expectancy communication.

It was reasoned that, if E bias were mediated by cueing alone, no differences between the feedback conditions would be evident in terms of the mean number of photos rated "success". In all three instances, E would be covertly saying to his Ss, "Give me a success response". Even in the No Feedback condition, where E was unaware of Ss' responses, cueing behavior by E would communicate to Ss the response expected of them. If the bias were mediated via reinforcement alone, differences in success
ratings would be expected between the groups. In the Correct Feedback condition, for example, E would be expected to positively reinforce ratings of "success" and negatively reinforce ratings of "failure". On the other hand, Es in the Reversed Feedback condition would be expected to positively reinforce Ss' "failure" responses and negatively reinforce their "success" responses. Consequently, an increase in the magnitude of "success" responses would be expected in the Correct Feedback condition, no change would be expected in the No Feedback condition, while a decrease in magnitude of "success" responses would be expected in the Reversed Feedback condition.

The findings of this study revealed no differences in the mean number of "success" responses between feedback conditions. This result was in line with the cueing hypothesis, but was incompatible with the reinforcement hypothesis. In both of the experimental groups (Correct Feedback and Reversed Feedback conditions), an increase was found in the mean number of "success" responses from the first ten to the last ten photos rated. The increase for the Correct Feedback condition approached significance ($p < .06$, one tail, $t = 1.73$, $df = 19$). The increase for the Reverse Feedback condition was not as great ($p < .20$, one tail, $t = 0.92$, $df = 19$).

The finding of no differences between feedback conditions along with the concomitant increase in the mean number of "success" responses as a function of trials, suggests that E expectancy may have been mediated by a cueing mechanism. The smaller increase in the number of "success" responses for the Reversed Feedback condition suggests that some reinforcement effects may have contributed to the bias communication. This would tend to
facilitate an increase in the number of expected responses in the Correct Feedback condition, since cueing and reinforcement would produce complementary effects. In the Reversed Feedback condition, however, cueing and reinforcement would tend to operate in disharmony and result in a smaller increase in the criterion measure. If the results are interpreted in this way, it is evident that the effects of cueing are much more profound than those of reinforcement for the photo-rating task.

It will be recalled that predata interaction between E and Ss was restricted in this study in an attempt to eliminate cueing effects during that phase of the experiment. A further restriction during the data collecting phase eliminated verbal communication between E and S. The implications are both disturbing and rewarding; disturbing because of the magnitude of E bias even when severe limitations were placed on E-S communication, rewarding because it reinforces the efforts being made to identify the specific cues involved in expectancy communication. The implications for studies of verbal conditioning are noteworthy. As it was pointed out earlier, the hypotheses of cueing and reinforcement are complementary in studies of verbal conditioning yet increases in the number of "correct" responses are usually attributed to effects of reinforcement. It is important that these studies be examined in greater detail. The possibility is very real in the light of this preliminary study that E may produce a greater effect by cueing the Ss as to how they should respond rather than by reinforcing them verbally for the correct response.

The present study is a follow-up of the study just described. Within the person perception paradigm it will explore the influence of verbal and
nonverbal communication in the mediation of E expectancy. Moreover, it
will attempt to determine if Ss' responses will be affected by different S
to E feedback conditions.

The hypotheses to be tested and their rationales are:

1. That Es given the same expectancy, success, but receiving
correct feedback, reversed feedback or no feedback will obtain similar
results on a photo-rating task. This hypothesis follows from the discussion
of the operation of cueing and reinforcement in the communication of E bias.
The Es who receive no feedback from their Ss would be unaware of whether or
not the Ss were responding in accordance with their hypothesis. In order
for a reinforcement mechanism to operate, Es must be aware of the Ss' responses. On the other hand, if Es are cueing their Ss as to how they
should respond, a bias effect may be in evidence. The Es who receive the
Ss' true ratings (correct feedback) would be expected to bias Ss in the
direction of their hypotheses. The same would be true whether the biasing
mechanism was a cueing or a reinforcing mechanism. For Es receiving
information exactly opposite to Ss' true ratings (reversed feedback), two
possibilities exist. (a) If cueing is the major biasing mechanism, Ss
would be covertly influenced by their Es to give success responses, con-
sequently a predominance of success responses would be obtained. (b) If
reinforcement is the major biasing mechanism, E would be expected to in-
advertently bias the Ss to give true ratings which disconfirm E's hypothesis
and, as a consequence, relatively fewer success responses would be obtained.

2. That in conditions in which verbal and nonverbal cues are
available more E bias effect will occur than in conditions in which only
nonverbal cues are permitted. This follows from studies which have shown that both verbal and nonverbal cues contribute to the bias effect.

Confirmation of the first hypothesis would provide support for an underlying cueing mechanism in F bias. Failure to confirm the first hypothesis would be consistent with an operant conditioning mechanism. Confirmation or disconfirmation of the second hypothesis has implications for the relative importance of verbal and nonverbal cues for expectancy communication.
CHAPTER III

METHOD

Experimenters and Subjects. Ten males enrolled in an advanced undergraduate psychology course at the University of British Columbia served as Es. All Es volunteered and were paid $3.00 for their part in the experiment. None of these students had prior experience with the person perception task used in the present study.

Subjects were 120 females, between 19 and 30 years old, enrolled in undergraduate education courses at U.B.C. All Ss were volunteers. They, too, had no prior experience with the stimulus materials.

Materials. The 20 standardized neutral photographs of faces that were used by Rosenthal (1963c) in his studies of experimenter bias in a person perception context were employed in the present study. The method by which these photographs were standardized by Rosenthal is described in Appendix A.

Feedback of responses to the photos in the nonverbal conditions was visually communicated to E by a display panel with an eight-point rating scale. The points on the scale were fixed with lights which when illuminated indicated the Ss' rating of the photo to E. On the other side of the panel, not visible to E, a similar eight-point rating scale was used by Ss in making their ratings. The points on the Ss' scale were fixed with push buttons so that by pressing a button on the scale, Ss made their ratings. A second panel of similar dimensions but without lights or buttons was used by E to obscure his data sheets from the Ss' view.

The Ss' push button rating scale was designed so that its poles were
reversible. That is, in the Correct Feedback condition a rating of 1 by Ss on the push button scale meant that Ss rated the photo "extreme failure" but in the Reversed Feedback condition, 1 meant "extreme success" (see Fig. 1). The rating scale on the front of the Ss' panel, visible only to E, was not reversible so that a rating of 1 always meant extreme failure to E (see Fig. 2). Thus, each rating that Ss made in the Reversed Feedback condition was unwittingly misinterpreted by E to mean exactly the opposite of what the Ss intended.

**Design.** The experiment was designed as a 2 x 3 x 10 factorial. The first independent variable was at two levels: Verbal vs. Nonverbal Communication (between E and Ss prior to and during the photo-rating task). The second independent variable was at three levels: Type of Feedback from Ss to E (i.e., correct feedback of responses, reversed feedback of responses, and no feedback of responses). The third (blocking) variable was at ten levels: the 10 Es in the study. The 2 x 3 x 10 factorial requires the 60 experimental groups shown in Table 1. Each group was constituted by two Ss.

Ten Es each ran 12 Ss in the photo-rating task. Half of the Ss (Groups 1 to 30) were assigned to the Verbal Communication condition. In this condition, there was no restriction on communication between E and Ss. The remaining Ss (Groups 31 to 60) were assigned to the Nonverbal Communication condition in which E and Ss were instructed not to talk to each other prior to or during the presentation of the photos. Within each of these conditions Ss were randomly assigned to one of three feedback conditions. Thus, each E ran two Ss from each of six treatment conditions.
Figure 1. Eight Point Success-Failure Rating Scale for the Correct Feedback Condition as Seen by Both E and S.

<table>
<thead>
<tr>
<th>FAILURE</th>
<th>SUCCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Eight Point Success-Failure Rating Scale for the Reversed Feedback Condition. A, as Seen by E; B, as Seen by S.

<table>
<thead>
<tr>
<th>FAILURE</th>
<th>SUCCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SUCCESS</th>
<th>FAILURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8</td>
<td></td>
</tr>
</tbody>
</table>
TABLE 1

EXPERIMENTAL DESIGN: 2 x 3 x 10 FACTORIAL

<table>
<thead>
<tr>
<th>Verbal Communication</th>
<th>Nonverbal Communication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Feedback (VC)</td>
<td>Correct Feedback (NVC)</td>
</tr>
<tr>
<td>Reverse Feedback (VR)</td>
<td>Reverse Feedback (NVR)</td>
</tr>
<tr>
<td>No Feedback (VK)</td>
<td>No Feedback (NVK)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experimenter 1</th>
<th>Group 1</th>
<th>Group 11</th>
<th>Group 21</th>
<th>Group 31</th>
<th>Group 41</th>
<th>Group 51</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimenter 2</td>
<td>Group 2</td>
<td>Group 12</td>
<td>Group 22</td>
<td>Group 32</td>
<td>Group 42</td>
<td>Group 52</td>
</tr>
<tr>
<td>Experimenter 3</td>
<td>Group 3</td>
<td>Group 13</td>
<td>Group 23</td>
<td>Group 33</td>
<td>Group 43</td>
<td>Group 53</td>
</tr>
<tr>
<td>Experimenter 4</td>
<td>Group 4</td>
<td>Group 14</td>
<td>Group 24</td>
<td>Group 34</td>
<td>Group 44</td>
<td>Group 54</td>
</tr>
<tr>
<td>Experimenter 5</td>
<td>Group 5</td>
<td>Group 15</td>
<td>Group 25</td>
<td>Group 35</td>
<td>Group 45</td>
<td>Group 55</td>
</tr>
<tr>
<td>Experimenter 6</td>
<td>Group 6</td>
<td>Group 16</td>
<td>Group 26</td>
<td>Group 36</td>
<td>Group 46</td>
<td>Group 56</td>
</tr>
<tr>
<td>Experimenter 7</td>
<td>Group 7</td>
<td>Group 17</td>
<td>Group 27</td>
<td>Group 37</td>
<td>Group 47</td>
<td>Group 57</td>
</tr>
<tr>
<td>Experimenter 8</td>
<td>Group 8</td>
<td>Group 18</td>
<td>Group 28</td>
<td>Group 38</td>
<td>Group 48</td>
<td>Group 58</td>
</tr>
<tr>
<td>Experimenter 9</td>
<td>Group 9</td>
<td>Group 19</td>
<td>Group 29</td>
<td>Group 39</td>
<td>Group 49</td>
<td>Group 59</td>
</tr>
<tr>
<td>Experimenter 10</td>
<td>Group 10</td>
<td>Group 20</td>
<td>Group 30</td>
<td>Group 40</td>
<td>Group 50</td>
<td>Group 60</td>
</tr>
</tbody>
</table>

as follows. Group VC, verbal correct feedback: the Ss in this group gave their true responses, verbatim, to E. (See instructions Appendix B.)

Group VR, verbal reversed feedback: this group of Ss gave a verbal rating interpreted by E as being opposite to the intended rating. It should be noted that Ss were instructed to give only the number of their rating. Any other rating by Ss such as saying "mild success" or "extreme failure", would not communicate the reverse feedback to E, but would in fact communicate the Ss' true ratings to E. (See instructions Appendix C.) Group VK, verbal no feedback: the Ss in this group indicated their responses by pressing the button corresponding to their rating after which they verbally communicated to E that they had made their rating. (The ratings were automatically recorded.) The ratings of this group of Ss were not revealed to E. (See
instructions Appendix D.) Group NVC, nonverbal correct feedback: the true responses by Ss in this group were shown on the Ss' panel. The Ss indicated their responses by pressing the button corresponding to their rating on the eight-point scale. The E recorded the responses as shown on the front of the Ss' panel. (See instructions Appendix E.) Group NVR, nonverbal reversed feedback: the Ss in this group rated the photos by pressing a button corresponding to their rating as was required of Group NVC. This time, however, the light on the Ss' panel corresponding to the rating opposite to the Ss' intended rating was shown, thus communicating reversed feedback to E, who recorded the ratings. (See instructions Appendix F.) Group NVK, nonverbal no feedback: (see instructions Appendix G). The responses of Ss in this condition were again indicated by Ss' pressing a button on the rating scale. In this condition, all of the lights on the display panel were activated irrespective of the particular button pressed by Ss. (The responses of the Ss were automatically recorded.) The E merely recorded on his data sheet that Ss had made a response.

The order of presentation of conditions for each E was randomly determined.

Procedure. In an attempt to prevent any E bias effects due to either the Principal Investigator (PI) or to the Es before the administration of the photo-rating task, interaction with Ss during this period was restricted. Appointments were arranged with Ss as to the time and place of the experiment prior to their assignment to any experimental condition. All other communications to Ss were in the form of written instructions.

The experiment was conducted in ten sessions. All experimental
conditions were represented in each session. When E arrived for the experi-
ment, he was seated at a table in the experimental room. In front of him,
on the table, was a panel (9" x 24") used to conceal his data sheets. Also
on the table, directly opposite E, was the display panel used by Ss to
indicate their photo ratings in the nonverbal experimental conditions. The
E was given a set of written instructions describing the experimental task
and the type of responses he should expect from his Ss. The reason advanced
for E's participation in the study stressed the need for replication by
researchers in order to expand the generality of the findings.

**Instructions to Experimenters.** You have been asked to
participate in a research project developing a test of
empathy. The reason for your participation in this
project is to standardize results of experiments of this
type. There is the problem in psychological research of
different examiners getting somewhat different data on
the same tests as a function of individual differences.
Therefore, to standardize the tests it is better
methodological procedure to use groups of experimenters.

You will now be asked to run a series of Ss and
to obtain from each, ratings of photographs. The ex-
perimenal procedure has been typed out for you and is
self-explanatory.

According to the preceding research of this nature,
some individuals have been found to possess greater
ability than others in discriminating between success and
failure from photographs of people. University students
tend to judge the photos in this experiment as "success
pictures". The type of Ss that you will be using have
averaged a 6.5 rating. Therefore, the Ss you are running
should average about a 6.5 rating.

Different methods of recording the Ss' responses
are used in this study to serve as a control for a
second part of the experiment to be carried out in a
later study.

**Experimental procedure for Experimenters.** In front of
you, you will find a sheet of paper for recording each
S's rating for each photo and a set of 20 numbered
pictures. Shuffle the photos and record the order in
which you will be presenting them before each S comes in.

When the S enters the room she will sit down opposite you and wait for you to present the first photo. Take the first photo and hold it in front of the S for 5 seconds. Some Ss will indicate their ratings verbally while others will press a button which will activate the light under the number corresponding to her rating. You will then check off the S's response on the recording sheet. Continue this procedure for the 20 photos. Do not let the S see any photo for longer than 5 seconds. If she does not respond within that time, wait for her response before presenting the next photo.

You will run 12 Ss. Continue the same procedure for each of your 12 Ss. (That is: (1) shuffle the photos, (2) record their order on the DATA SHEET, (3) wait for the next S to come in.)

Four of your Ss' responses will be indicated by all of the lights going on at the same time. For these Ss, merely show that the photo has been responded to by making a check mark opposite the trial number.

Six of your Ss will have been instructed to say nothing prior to or during the photo presentations. I will tell you what Ss have been so instructed. It is important that you say nothing to these Ss. If for any reason you should say anything, please write down the exact words used and the situation which forced you to say them.

After reading the instructions, E was questioned about the purpose of the experiment, the procedure to be used, and the type of responses he should expect to obtain. If E was not clear on any one of these points, he was told to read the instructions again to insure that he had complete comprehension of the situation. When it was apparent that E understood his role, he was left alone in the experimental room to prepare for his first S.

When Ss arrived for the experiment, they entered an anteroom off the experimental room. (Only one S was admitted at a time.) A large sign on the door of the adjacent experimental room instructed S to sit down and
read the typed sheet of instructions that she would find on the chair (see Appendix B, C, D, E, F, and G). After reading the instructions, Ss were instructed to wait until signaled to go into the experimental room. (The signal was a light remotely controlled by the Principal Investigator.) On entering the experimental room, S sat opposite E, behind the push button rating scale. When she was seated, E presented the first photo.

All E-S interactions were monitored in an observation room by the Principal Investigator. The experimental room was equipped with concealed microphones and a concealed TV camera so that neither E nor Ss were aware of being observed. The observation permitted PI to note any deviations from the standard experimental procedure. More important, it enabled PI to know when to modify the Ss' rating scale and when to signal for Ss to enter the experimental room.
CHAPTER IV
RESULTS

Table II shows the Ss' mean photo ratings in relation to the type of feedback permitted in the experimental situation. A brief inspection of the overall data reveals only small variations among the mean ratings. The means cluster around the 4.50 midpoint of the rating scale and range between 3.70 and 5.38.

TABLE II
EXPERIMENTERS' MEAN PHOTO RATING SCORES FOR THE SIX EXPERIMENTAL CONDITIONS

<table>
<thead>
<tr>
<th>Experimenter</th>
<th>Verbal Communication</th>
<th>Nonverbal Communication</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct Feedback</td>
<td>Reverse Feedback</td>
</tr>
<tr>
<td></td>
<td>VC</td>
<td>VR</td>
</tr>
<tr>
<td></td>
<td>Correct Feedback</td>
<td>Reverse Feedback</td>
</tr>
<tr>
<td></td>
<td>NVC</td>
<td>NVR</td>
</tr>
<tr>
<td>1</td>
<td>4.83</td>
<td>4.60</td>
</tr>
<tr>
<td>2</td>
<td>3.75</td>
<td>4.15</td>
</tr>
<tr>
<td>3</td>
<td>4.53</td>
<td>4.35</td>
</tr>
<tr>
<td>4</td>
<td>4.58</td>
<td>3.94</td>
</tr>
<tr>
<td>5</td>
<td>4.98</td>
<td>4.63</td>
</tr>
<tr>
<td>6</td>
<td>5.38</td>
<td>4.65</td>
</tr>
<tr>
<td>7</td>
<td>4.03</td>
<td>3.98</td>
</tr>
<tr>
<td>8</td>
<td>4.45</td>
<td>4.35</td>
</tr>
<tr>
<td>9</td>
<td>4.08</td>
<td>4.70</td>
</tr>
<tr>
<td>10</td>
<td>4.43</td>
<td>4.85</td>
</tr>
<tr>
<td>Mean Rating</td>
<td>4.50</td>
<td>4.42</td>
</tr>
</tbody>
</table>

1The ratings used in this and other analyses were the Ss' actual ratings and not those recorded by Es.
For the purpose of analysis each S's data were divided into four blocks of five trials each. A summary of the Trial Block means in relation to treatment conditions is presented in Table III.

### Table III

#### MEAN PHOTO RATINGS OVER BLOCKS OF TRIALS FOR EACH OF THE SIX TREATMENT CONDITIONS

<table>
<thead>
<tr>
<th>Treatment Condition</th>
<th>Trial Block</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Correct Feedback</td>
<td>4.41</td>
<td>4.82</td>
<td>4.19</td>
<td>4.58</td>
</tr>
<tr>
<td>Verbal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reversed Feedback</td>
<td>4.25</td>
<td>4.43</td>
<td>4.55</td>
<td>4.44</td>
</tr>
<tr>
<td>No Feedback</td>
<td>4.21</td>
<td>4.19</td>
<td>4.59</td>
<td>4.48</td>
</tr>
<tr>
<td>Correct Feedback</td>
<td>4.27</td>
<td>4.60</td>
<td>4.40</td>
<td>4.28</td>
</tr>
<tr>
<td>Nonverbal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reversed Feedback</td>
<td>4.44</td>
<td>4.98</td>
<td>4.77</td>
<td>4.99</td>
</tr>
<tr>
<td>No Feedback</td>
<td>4.11</td>
<td>4.61</td>
<td>4.41</td>
<td>5.04</td>
</tr>
<tr>
<td>TOTALS</td>
<td>4.28</td>
<td>4.61</td>
<td>4.49</td>
<td>4.64</td>
</tr>
</tbody>
</table>

The data were analysed using a four factor repeated measures ANOVA. Table IV summarizes the results of the analysis. Only one factor was found to have a significant effect, that for Trial Blocks ($F = 5.25, df = 3, 321, p < .01$). The means averaged over Es and treatments for the four trial blocks were, in order, 4.28, 4.61, 4.49, and 4.64. Figure 3 shows these means as a function of trial blocks. Inspection of this figure reveals a relatively large increase in the magnitude of mean ratings from the first to the second block of trials, a slight drop from the second to the third block.
and an increase from the third to the fourth block.

TABLE IV
SUMMARY OF THE FOUR FACTOR REPEATED MEASURES ANOVA FOR THE 20 TRIAL PHOTO-RATING TASK

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal (V)</td>
<td>1</td>
<td>2.591</td>
<td>2.591</td>
<td>1.95</td>
</tr>
<tr>
<td>Feedback (FD)</td>
<td>2</td>
<td>2.648</td>
<td>1.324</td>
<td>1.00</td>
</tr>
<tr>
<td>Experimenter (E)</td>
<td>9</td>
<td>15.825</td>
<td>1.758</td>
<td>1.32</td>
</tr>
<tr>
<td>VXFD</td>
<td>2</td>
<td>4.844</td>
<td>2.422</td>
<td>1.82</td>
</tr>
<tr>
<td>VXE</td>
<td>9</td>
<td>7.541</td>
<td>0.838</td>
<td>0.63</td>
</tr>
<tr>
<td>EXFD</td>
<td>18</td>
<td>19.400</td>
<td>1.078</td>
<td>0.81</td>
</tr>
<tr>
<td>VXEXFD</td>
<td>18</td>
<td>22.508</td>
<td>1.251</td>
<td>0.94</td>
</tr>
<tr>
<td>Error (Between Ss)</td>
<td>60</td>
<td>79.759</td>
<td>1.329</td>
<td></td>
</tr>
<tr>
<td>Trial Blocks (TR)</td>
<td>3</td>
<td>9.271</td>
<td>3.090</td>
<td>5.25**</td>
</tr>
<tr>
<td>TRXV</td>
<td>3</td>
<td>1.689</td>
<td>0.563</td>
<td>0.96</td>
</tr>
<tr>
<td>TRXFD</td>
<td>6</td>
<td>6.006</td>
<td>1.001</td>
<td>1.70</td>
</tr>
<tr>
<td>TRXE</td>
<td>27</td>
<td>15.562</td>
<td>0.576</td>
<td>0.98</td>
</tr>
<tr>
<td>Error (Within Ss)</td>
<td>321</td>
<td>189.080</td>
<td>0.589</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>479</td>
<td>376.730</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**significant at the 1% level.

A trend analysis for the Trial Blocks factor is summarized in Table V. The F tests for both linear and cubic trends exceeded the value required for the .01 level of significance. Thus, the relation between Trial Blocks and mean photo-ratings includes both linear and cubic components, providing
Figure 3. Grand mean photo ratings obtained for the four trial blocks.
statistical confirmation for the trends visually manifested in Figure 3.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>3</td>
<td>3.090</td>
<td>5.25</td>
</tr>
<tr>
<td>Linear</td>
<td>1</td>
<td>5.299</td>
<td>9.00**</td>
</tr>
<tr>
<td>Quadratic</td>
<td>1</td>
<td>0.910</td>
<td>1.55</td>
</tr>
<tr>
<td>Cubic</td>
<td>1</td>
<td>3.052</td>
<td>5.18**</td>
</tr>
<tr>
<td>Error</td>
<td>321</td>
<td>0.589</td>
<td></td>
</tr>
</tbody>
</table>

**significant at the 1% level.

To determine if the means averaged over Es and treatments increased significantly as a function of the number of trials employed, comparisons among the block means were made using the method of orthogonal polynomials (Hays, 1965, pp. 466-471). A cumulative effect on the magnitude of Ss' ratings was predicted from the three processes described as possible mediators of E expectancy (i.e., operant conditioning of Ss' responses, operant conditioning of Es' communicative behavior, and cueing). It was important, therefore, to determine if a cumulative effect did occur, and if it occurred, to determine how quickly it was manifested. The comparisons chosen to answer these questions were: the average of Blocks 1 and 2 vs. the average of Blocks 3 and 4, and Block 1 vs. Block 2. A summary of the comparisons is presented in Table VI. The comparisons showed the average rating of Blocks 3 and 4 to be significantly greater than the average of
Blocks 1 and 2 ($p < .001$). Moreover, when Blocks 1 and 2 were compared, the mean of Block 2 was greater ($p < .001$). Thus, the increase in ratings over trial blocks was significant and was manifested by the end of the second block of trials.

**TABLE VI**

COMPARISONS OF THE GRAND MEANS FOR THE FOUR TRIAL BLOCKS

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>$df$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocks 1 &amp; 2 vs. 1 &amp; 2</td>
<td>116</td>
<td>10.32****</td>
</tr>
<tr>
<td>Blocks 3 &amp; 4</td>
<td>116</td>
<td>5.27****</td>
</tr>
<tr>
<td>Block 1 vs. Block 2</td>
<td>116</td>
<td>5.27****</td>
</tr>
</tbody>
</table>

****significant at the 0.1% level.

The results of these tests indicate that $S$s gave ratings of increasing magnitude as the number of trials increased, irrespective of the type of experimental condition to which they were assigned. Apparently, neither the mode of communication nor the type of $E-S$ feedback affects the magnitude of the photo-ratings reliably but the number of trials does make a substantial difference. The data are consistent with the initial hypothesis which predicted no difference between the three feedback conditions.

The second hypothesis, that verbal and nonverbal cues would produce a greater $E$ expectancy effect than would nonverbal cues alone is not supported by the results of the four factor ANOVA. To gain a greater appreciation of
the effect of the restriction of verbal cues, separate ANOVAs for the ratings within each of the feedback conditions were carried out.

The separate ANOVAs for the ratings within each of the feedback conditions showed a greater increase in ratings under those conditions in which verbal cues were restricted. In Group NVR, for example, a two factor repeated measures ANOVA showed a significant main effect due to the Trial Block factor \( (F = 3.81, df = 3,27, p < .05) \). The results of the ANOVA for Group NVR are summarized in Table VII. The trial block means were 4.44, 4.98, 4.77, and 4.99 for Blocks 1 to 4 respectively, indicating an increase in mean photo-ratings as the number of trials increase. A similar finding was evident from the results of a two factor repeated measures ANOVA for Group NVK, summarized in Table VIII. This analysis showed a significant effect due to an increase in the mean photo-ratings over blocks of trials \( (F = 7.25, df = 3,27, p < .01) \). The means for Trial Blocks 1 to 4 were 4.11, 4.61, 4.41, and 5.04. The pattern established in Groups NVR and NVK

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimenter (E)</td>
<td>9</td>
<td>0.549</td>
<td>0.61</td>
</tr>
<tr>
<td>Trials (TR)</td>
<td>3</td>
<td>1.326</td>
<td>3.81*</td>
</tr>
<tr>
<td>TRXE</td>
<td>27</td>
<td>0.348</td>
<td>0.39</td>
</tr>
<tr>
<td>Within</td>
<td>40</td>
<td>0.894</td>
<td></td>
</tr>
</tbody>
</table>

*p < .05
TABLE VIII

ANALYSIS OF VARIANCE OF OBTAINED PHOTO RATINGS FOR GROUP NVK

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimenter (E)</td>
<td>9</td>
<td>0.901</td>
<td>1.35</td>
</tr>
<tr>
<td>Trials (TR)</td>
<td>3</td>
<td>3.045</td>
<td>7.25**</td>
</tr>
<tr>
<td>TRxE</td>
<td>27</td>
<td>0.420</td>
<td>0.63</td>
</tr>
<tr>
<td>Within</td>
<td>40</td>
<td>0.670</td>
<td></td>
</tr>
</tbody>
</table>

**p < .01

was not upheld, however, in Group NVC. The two factor ANOVA for this group failed to show any significant effects (see Table IX). Similar ANOVAs for each of the three Verbal Feedback conditions showed no Trial Block effects (see Tables X, XI, and XII). Contrary to the second hypothesis then, there was a more pronounced increase in mean photo-ratings over blocks of trials in the Nonverbal treatment conditions than in the Verbal conditions. It should be noted, however, that the failure of the four factor ANOVA to show
a significant interaction between the Trial Block and Verbal Factors demonstrates that no statistically significant difference exists between the Verbal and Nonverbal conditions with respect to the $S_s$' mean photo-ratings.

### TABLE X

**ANALYSIS OF VARIANCE OF OBTAINED PHOTO RATINGS FOR GROUP VC**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimenter</td>
<td>9</td>
<td>1.849</td>
<td>1.66</td>
</tr>
<tr>
<td>Trials</td>
<td>3</td>
<td>1.420</td>
<td>1.72</td>
</tr>
<tr>
<td>TRXE</td>
<td>27</td>
<td>0.824</td>
<td>0.74</td>
</tr>
<tr>
<td>Within</td>
<td>40</td>
<td>1.113</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE XI

**ANALYSIS OF VARIANCE OF OBTAINED PHOTO RATINGS FOR GROUP VR**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimenter</td>
<td>9</td>
<td>0.815</td>
<td>1.69</td>
</tr>
<tr>
<td>Trials</td>
<td>3</td>
<td>0.309</td>
<td>0.62</td>
</tr>
<tr>
<td>TRXE</td>
<td>27</td>
<td>0.501</td>
<td>1.04</td>
</tr>
<tr>
<td>Within</td>
<td>40</td>
<td>0.482</td>
<td></td>
</tr>
</tbody>
</table>
TABLE XII

ANALYSIS OF VARIANCE OF OBTAINED PHOTO RATINGS FOR GROUP VK

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimenter (E)</td>
<td>9</td>
<td>2.365</td>
<td>2.70*</td>
</tr>
<tr>
<td>Trials (TR)</td>
<td>3</td>
<td>0.790</td>
<td>1.11</td>
</tr>
<tr>
<td>TRXE</td>
<td>27</td>
<td>0.711</td>
<td>0.81</td>
</tr>
<tr>
<td>Within</td>
<td>40</td>
<td>0.876</td>
<td></td>
</tr>
</tbody>
</table>

*p < .025

**Supplementary Analysis.** In addition to the above analysis using the 20 trials of the photo-rating task it seemed of value to subject to closer scrutiny the first ten trials only. The reason for this decision stems from the fact that the majority of studies employing the photo-rating task have used only ten trials. In order to compare the findings of the present study with earlier findings, a four factor repeated measures ANOVA was performed on the data from the first ten trials. The ten trials were considered as two blocks of five trials each.

Table XIII summarizes the results of the ANOVA. Only the F value for the Trial Block factor exceeded the F required for statistical significance (F = 10.89, df = 1, 107, p < .005). This finding supports the results of the 20 trial ANOVA.

A trend analysis over the grand means for the first 9 trials shows a significant linear trend (F = 13.46, df = 1, 856, p < .005). The analysis is summarized in Table XIV. Figure 4 shows the relationship between mean ratings and trials (a straight line is plotted by the method of least squares for the first ten trials).
<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal (V)</td>
<td>1</td>
<td>0.819</td>
<td>0.819</td>
<td>1.09</td>
</tr>
<tr>
<td>Feedback (FD)</td>
<td>2</td>
<td>3.204</td>
<td>1.602</td>
<td>2.13</td>
</tr>
<tr>
<td>Experimenter (E)</td>
<td>9</td>
<td>5.410</td>
<td>0.601</td>
<td>0.80</td>
</tr>
<tr>
<td>VXFD</td>
<td>2</td>
<td>3.080</td>
<td>1.540</td>
<td>2.05</td>
</tr>
<tr>
<td>VXE</td>
<td>9</td>
<td>2.384</td>
<td>0.265</td>
<td>0.35</td>
</tr>
<tr>
<td>EXFD</td>
<td>18</td>
<td>17.849</td>
<td>0.992</td>
<td>1.32</td>
</tr>
<tr>
<td>VXEXFD</td>
<td>18</td>
<td>14.307</td>
<td>0.795</td>
<td>1.06</td>
</tr>
<tr>
<td>Error (Between)</td>
<td>60</td>
<td>45.140</td>
<td>0.752</td>
<td></td>
</tr>
<tr>
<td>Trials (TR)</td>
<td>1</td>
<td>6.275</td>
<td>6.275</td>
<td>10.89***</td>
</tr>
<tr>
<td>TRXV</td>
<td>1</td>
<td>1.065</td>
<td>1.065</td>
<td>1.85</td>
</tr>
<tr>
<td>TRXFD</td>
<td>2</td>
<td>0.208</td>
<td>0.104</td>
<td>0.18</td>
</tr>
<tr>
<td>TRXE</td>
<td>9</td>
<td>3.982</td>
<td>0.442</td>
<td>0.77</td>
</tr>
<tr>
<td>Error (Within)</td>
<td>107</td>
<td>61.669</td>
<td>0.576</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>239</td>
<td>165.390</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

***Significant at the 0.5% level
In sum, the analysis of the first ten trials leads to the prediction of a significantly increasing linear relationship between the mean photo-ratings and the trial factor for the population of Es represented by the sample employed in this study. Thus, the conclusions reached using the first half of the data parallel those reached using all 20 trials with one exception. The relationship between ratings and trials is considered to be linear for the case of 10 trials, whereas, for the case of 20 trials the relationship contains both linear and cubic components.
Figure 4. Grand mean photo ratings obtained for the twenty trials.
It will be recalled that rather severe restrictions on communication between Es and Ss were invoked in the present study. They included written instructions for Ss, and the elimination of any Principal Investigator and E contact with Ss prior to data collection. The restrictions were an attempt to insure that E expectancies were not communicated prior to the administration of the photo-rating task. These predata restrictions may account for the small range of mean ratings obtained in the present study. They may also account for the failure of any mean to reach the 6.5 average that Es had been told to expect. It should be noted, however, that the range of judgements and the magnitude of E bias effect in the present study is in keeping with other studies of E bias utilizing this task (Fode, 1965; Rosenthal and Fode, 1963b; Rosenthal, Persinger, Vikan-Kline and Fode, 1963b).

The Rosenthal et al. (1963b) study investigated the influence of early data returns on E expectancy effects. It is particularly relevant to this discussion because of its methodological similarity to the present study. Both studies employed 20 photos in the standard photo-rating task and all Es were led to expect ratings 50% higher than the neutral point on the rating scale (i.e., +5 in the case of Rosenthal's study and 6.5 in the present one).  

1: The Rosenthal et al. (1963b) study employed a 20 point success-failure rating scale running from -10 (extreme failure) to +10 (extreme success) with midpoint at 0. The eight point rating scale used in the present study ran from 1 (extreme failure) to 8 (extreme success) with its midpoint at 4.5. Comparisons between the two scales were made by multiplying deviations from the 4.5 midpoint of the eight point scale by a factor of 2.5. Thus, a rating of 6.5 on the eight point scale is equivalent to 5 on the 20 point scale.
In the Rosenthal et. al. study the greatest mean rating obtained by any 
E from his naive test Ss was 1.19 compared to the lowest rating of -0.09. 
This represents a range of 1.28 which is less than the range of 1.68 
obtained in the present study.

An interesting ancillary finding of the Rosenthal et. al. (1963b) 
study was that none of the pairs of accomplices actually gave ratings of 
+5 or -5 as they had been instructed to do in confirming or disconfirming 
the hypotheses of their Es. Their mean ratings ranged from -4.38 to 3.68. 
It is not surprising then, that the mean ratings given by naive Ss in the 
present study failed to reach the 6.5 average that Es were told to expect. 
If accomplices will not give extreme ratings when explicitly told to do so 
by the investigator, then it is not surprising that even a less extreme 
performance is obtained from naive Ss in response to the covert cues of 
their Es.

Three mechanisms were considered possible mediators of E expectancy: 
operant conditioning of Ss' responses, operant conditioning of E's 
communicative behavior, and cueing of Ss as to how they are expected to 
respond. Each of these mechanisms will be discussed in light of the 
findings of the present study.

If operant conditioning of Ss' responses were the mediator of Es' 
expectancy then it was predicted that Ss would be differentially reinforced 
for expected and unexpected responses. Thus, Ss giving a "success" 
response would be positively reinforced in the Correct Feedback condition, 
given no reinforcement in the No Feedback condition, and negatively rein-
forced in the Reversed Feedback condition. As a result, differences among
the feedback conditions should be evident with respect to the magnitude of "success" ratings obtained by Es. Since no differences were found in the present study, it would appear that Ss were not being reinforced for their responses. This interpretation is consistent with the author's preliminary study (Moffat, 1966) which showed that the type of S to E feedback had no effect on the Ss' photo-ratings. Moreover, the study by Rosenthal, Fode, Vikan-Kline, and Persinger (1964) demonstrated that operant conditioning was neither a necessary condition nor a necessary augmenting factor in the operation of the E outcome-bias phenomenon; a finding consistent with the present study.

The suggestion made in Chapter II, that bias effects due to operant conditioning may have been obscured by biasing in the predata collecting phase of the Rosenthal et al. (1964) study, is not supported. Even when E-S interaction is eliminated prior to the data collecting phase of the experiment there is still no evidence for operant conditioning in the present study during actual data collection.

The second mechanism considered as a mediator of E expectancy was based on Rosenthal's suggestion that Ss may shape their Es' communicative behavior by a process of operant conditioning (Rosenthal, 1965). From this hypothesis, the shaping of Es' behavior depends upon the expectancy of Es and the type of feedback they obtain from their Ss. Specifically, a success response by Ss which confirms Es expectancies should positively reinforce any covert communicative behavior that immediately preceded Ss' response. An incorrect response, on the other hand, should negatively reinforce Es' communicative behavior that just preceded that response.
Evidence for the operation of the second mechanism was expected by comparing the mean photo-ratings of the three feedback conditions. (In other words, the first and second mechanism would have produced the same data.) Experimenters given correct feedback would have been expected to develop a communication pattern that influenced their Ss to respond "success". Conversely, Es given reversed feedback would have been expected to unwittingly influence their Ss to respond "failure", while Es in the No Feedback condition of course, could not have differentially communicated to their Ss. The largest mean rating was predicted for the Correct Feedback condition. The finding in the present study of no differences between feedback conditions with respect to mean photo-ratings suggests that Es' covert communicative behavior was not shaped by the reinforcing effects of their Ss' hypothesis confirming or disconfirming responses.

Cueing was the third process considered as a possible mediator of E expectancy. This process was described as being independent of any reinforcement process between Ss and Es. Therefore, the type of response Ss gave would have had no particular bearing on the type of cues that E communicated to his Ss. That is, whether or not Es' expectancies were confirmed would be irrelevant to Es' behavior. According to this interpretation, prior to each rating Es are, in effect, covertly saying to their Ss, "Give me a 'success' response". Consequently, even in the No Feedback condition Ss are "told" how to respond.

If cueing were a powerful mediator of E bias, its effects should be manifested in the ratings of the first photo. The only means of assessing a cueing effect at this stage would be to compare the ratings obtained by Es
with different expectancies. As all Es had the same expectancy, they would be expected to influence their Ss in the same manner. Consequently, E influence could not be examined at this stage of the task in this experiment.

Another way to assess the cueing effect is to examine the photo-ratings over trials. If cueing were a mediator of E bias we might expect a progressive increase in the magnitude of "success" responses for all treatment conditions, irrespective of the type of S feedback permitted. There are two reasons to expect a cumulative cueing effect. First, the anxiety experienced by both Es and Ss may interfere with the communication and interpretation of Es' cues. As the experiment progresses, however, anxiety may be dissipated as Es and Ss become acquainted with the task and with each other.

The second factor that may influence the cueing effect stems from the nature of the task. It seems likely that Ss at first would try to base their photo-ratings on the characteristics of the person in the photograph. Since the photos are selected for their lack of distinctive cues, Ss may find this approach extremely frustrating. It seems reasonable that Ss would begin to rely less on the meager information available from the photos and would concentrate more on situational and E cues, or what Orne (1962) calls the "demand characteristics" of the situation. Masling (1966) lends support to this line of reasoning in his discussion of behavior under conditions of ambiguity. He cites evidence to suggest that an ambiguous experimental situations is more conducive to biasing by Es. Moreover, under such conditions

---

2This suggestion was supported by informal comments of Ss after they had completed the task.
of ambiguity he feels that Ss will utilize all the available information in his search for meaning and guidance. Thus, the importance of minimal cues is magnified in situations of ambiguity. The additivity of these cues should result in an increase in the magnitude of "success" responses as a function of the number of trials employed.

The results of the present study are congruent with the cueing process interpretation of E(expectancy communication. The finding of no difference between feedback conditions with respect to mean photo-ratings leads to the conclusion that, if there were an E(expectancy effect, it is independent of the type of feedback that Es obtain from their Ss. This finding alone precludes the operation of a reinforcement mechanism. It does not, however, provide proof that cueing was in operation. Evidence for the cueing mechanism resides in the finding of an increase in the overall magnitude of mean photo-ratings as the number of trials increase.

When the data for each feedback condition is examined, only two of the feedback conditions showed a significant increase in ratings. However, each of the other four feedback conditions did show an increase in mean photo-ratings from the first to the fourth trial block. On the basis of this evidence and considering the discussion of the cueing process, it is concluded that the mechanism responsible for communicating Es' expectancies during the photo-rating period of the person perception task probably is cueing.

If the interpretation offered to account for the results is correct, then the present experiment in no way contradicts the findings of earlier studies. The interpretation must, however, be tempered in the light of an
alternative interpretation adopted from the person perception study of Rosenthal, Fode, Vikan-Kline and Persinger (1964), described earlier. It will be recalled that in their study half the Es were told to expect +5 ratings of success and the other half were told to expect -5 ratings of failure from their Ss on a photo-rating task. The Ss task was to rate a series of photographs on a twenty point scale ranging from +10 to -10 for success or failure. When the mean ratings for Es with opposite expectancies were plotted over the ten trials, both groups of Es obtained progressively higher mean ratings. These authors suggested that the increase in ratings may have been due to what they called a "photograph effect, a sequence effect or a photograph by sequence interaction effect" (Rosenthal et al., 1964, p. 74). This suggestion may be offered as an alternative interpretation of the increase in magnitude of mean photo-ratings over trials, as was found in the present study. It should be noted, however, that other studies using the photo-rating task did not produce the photo or sequence effect. To clarify the issue it would be informative to run another treatment group employing Es with no expectancy or with an expectancy opposite to the one used here. An increase in ratings as a function of trials in these groups would suggest that something other than the E's expectancy determines how Ss respond. Only another experiment can hope to resolve this issue.

The restrictions on verbal communication during the task had no effect on Ss' ratings. This finding is interpreted as evidence that verbal cues make no significant contribution to the communication of E expectancy. The interpretation is strengthened somewhat by the additional finding that
two out of three Nonverbal treatment conditions contained significant mean increases over trials whereas none of the conditions permitting verbal cues showed significant increases.

Rosenthal and Fode (1963b) showed that eliminating verbal cues reduced the bias effect, a finding at variance with Moffat (1966), and the present study. It was pointed out earlier that the discrepancy in the findings of these studies may be due to a confounding of instruction reading behavior with the elimination of verbal cues in the Rosenthal and Fode study. That is, not only did they eliminate verbal cues in the Nonverbal condition, they also eliminated Es' instruction reading behavior which may have resulted in the elimination of important visual-kinesic cues germane to the communication of E's expectancy. This criticism was circumvented in the present study by eliminating instruction reading behavior from both the Nonverbal conditions and the Verbal conditions. As a result, a true test of the effect of verbal as compared to nonverbal cues was made during data collection. The conclusion is that verbal cues are not necessary for communicating E expectancy during the task administration phase of the experiment. The argument does not however detract from the hypothesis that verbal cues can be important communicators of E bias (Friedman, Kurland and Rosenthal, 1965; Rosenthal, Fode, Vikan-Kline and Persinger, 1964; Weick, 1963).

The same conclusions are drawn in the present study whether the analyses are based on all the data or on the data of the first ten trials only. Perhaps this explains why the majority of experiments in E outcome-bias employ the ten trial photo-rating task. Additional trials seem to be
unwarranted because of the redundant information they provide.\(^3\)

The unexpected reversal in the trend of the ratings from Block 2 to Block 3 in the original analysis is difficult to explain. The simplest tentative explanation may be based on an examination of the phenomenology of the Ss in the experimental situation. The Ss may have been aware of the fact that they were giving ratings of increasing magnitude in the first ten or so trials. In order to compensate for the apparently unjustified increase, Ss may have intentionally reduced their ratings at this point only to increase them in latter trials as they were "cued" to do by their Es. Other investigators have termed this the "boomerang effect" (Silverman, 1965) or the "screw you effect" (Masling, 1966). The "screw you" motive may rise from Ss' curiosity as to what would happen if they varied their responses or it may be an attempt to show E that he cannot control their behavior.

There is no way to determine how Ss perceived their role in the present experiment. Certainly Ss are not passive responders, devoid of intelligence, purpose, or orienting ability. Unfortunately, that is how they are usually treated. Perhaps we overlook the most important source of data in the psychological experiment by not having Ss report their hypotheses of what the experiment was designed to test and how they responded the demand characteristics of the experimental situation. Masling may be correct in predicting that, "our Ss will teach us something about psychology" (Masling, 1966, p. 96).

\(^3\)An opinion shared by Rosenthal in a personal communication.
Previous studies have shown that Es may obtain data from their Ss that are in accordance with Es' experimentally induced or idiosyncratic expectations. It seems that Es unwittingly communicate their expectancies to Ss by using subtle gestures together with covert verbal and visual cues. Some observers of filmed E-S interactions believe that Es' expectancies may be communicated in the brief predata collecting phase of the experiment when Es greet their Ss and read the experimental instructions. Other observers believe that expectancy communication is mediated by an operant conditioning process and does not take place until after Ss begin responding. In other words, Es are believed to differentially reinforce the responses of their Ss. Experimental data bearing on the hypothesis of an operant conditioning process, as a mediator of E expectancy, have been inconclusive.

The present study examined the effect of S to E feedback, mediated by verbal and nonverbal channels of communication upon the E outcome bias phenomenon. Since previous studies have shown that expectancy communication may take place prior to the Ss' first response, it would seem that an operant conditioning process would not be necessary for the communication of E expectancy. Moreover, as the communication may take place before Ss begin to respond, it would seem that the process responsible for mediating expectancy communication is not dependent upon the Ss' responses or feedback. An alternative process was hypothesized as mediating E expectancy, the process of cueing. This process was hypothesized as being independent of S to E feedback.
Ten Es each ran 12 Ss on a photo-rating task in a study purporting to be a research project developing a test of empathy. The Ss were required to examine 20 standardized neutral photographs of faces and to rate each one on the degree of success or failure that the person pictured had been experiencing. The Es had been led to expect a predominance of success responses from their Ss.

The effect of different types of S to E feedback was examined by permitting correct, reversed, or no feedback of Ss' photo rating responses. The relationship between Ss' feedback and E bias effects was examined by comparing the grand mean photo ratings from each feedback condition, and by comparing the mean photo ratings over blocks of trials for each feedback condition. The outcome of these comparisons had relevance for the type of mediating process believed to be responsible for the communication of E expectancy. Three processes were discussed in the light of the present findings (operant conditioning of Ss' responses, operant conditioning of Es' communicative behavior, and cueing).

The effect of mode of E-S communication on E bias was examined by permitting or restricting verbal communication between E and Ss for each of the three feedback conditions. The relationship between mode of communication and E bias was determined by comparing the mean photo ratings of Ss in the verbal and nonverbal conditions.

Severe restrictions were placed upon E-S communication in the early predata collecting phase of the experiment in an attempt to eliminate bias effects at this stage. As bias effects were obtained in the present study, it was concluded that they must have occurred solely as a result of events in the later data collecting phase of the experiment. This conclusion,
together with the findings of other studies which have found significant E bias effects due to events in the predata collecting phase of the experiment, should reinforce the efforts of researchers who are making filmed analyses of both stages of E-S interactions in an attempt to identify the specific cues used in expectancy communication. Continuing research on E expectancy effects must identify these specific cues and determine how the cues are used in expectancy communication.

Some indication as to how cues are used in E expectancy communication is provided by the present study. Of the three processes considered as mediators of E expectancy, the process of cueing seems to afford the best explanation for the mediation of E expectancy effects. The concept of cueing is also consistent with the findings of other studies on E bias effects, but further research is needed to expand and strengthen the generality of the cueing concept.

Although previous studies have shown that restricting visual and verbal communication during the psychological experiment does reduce bias effects, the present study showed that there was no reduction in E bias effect when verbal cues were restricted. Future investigations should, therefore, be addressed to the problem of determining how much variance of the E expectancy effect is contributed by what mode of communication in what phase of the experiment. Only then can significant gains be made in developing E bias control techniques.

What are the implications of the present study for the development of E bias control techniques? The technique of minimizing E-S contact was used in the present study, but even when S to E feedback was eliminated or
reversed, no differential effect of influence was produced. It must be concluded from these results that minimized contact does not eliminate bias effect. Considering the methodological inconvenience of implementing minimized contact, the advantage of such a control measure is questionable. Other control techniques are discussed by Rosenthal (1966) and include training of Es, use of mechanical observation techniques (video tape and film), "blind" and minimized contact, automated data collection, and absence of from the experimental room. Some of these control measures have proven effective while others are still in the experimental stage. Future research must attempt to develop and improve these control measures.
BIBLIOGRAPHY


Lord, E. Experimentally induced variations in Rorschach performance. Psychological Monographs, 1950, 64 (10, Whole No. 316).


Pfungst, O. Clever Hans (the horse of Mr. von Osten): A contribution to experimental, animal, and human psychology. (Translated by C. L. Rahn.) New York: Holt, 1911.


Rosenthal, R., & Fode, K. L. Psychology of the scientist: V. Three experiments in experimenter bias. Psychological Reports, 1963, 12, 491-511. (b)


APPENDIX A

STANDARDIZATION PROCEDURE FOR THE PERSON PERCEPTION TASK

Fifty-seven photographs of faces ranging in size from 2 x 3 cm to 5 x 6 cm were cut from a weekly news magazine and mounted on 3 x 5 in. white cards. These were presented to 70 male and 34 female students enrolled in an introductory psychology class at the University of North Dakota. Subjects were instructed to rate each photo on a rating scale of success or failure. The scale shown in Figure 1 ran from -10, extreme failure, to +10, extreme success, with intermediate labeled points. Each subject was seen individually by the author who read to each the following instructions:

Figure 1

The Empathy Test Rating Scale

<table>
<thead>
<tr>
<th>Extreme Failure</th>
<th>Moderate Failure</th>
<th>Mild Failure</th>
<th>Mild Success</th>
<th>Moderate Success</th>
<th>Extreme Success</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10</td>
<td>-9</td>
<td>-8</td>
<td>-7</td>
<td>-6</td>
<td>-5</td>
</tr>
<tr>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>+1</td>
<td>+2</td>
</tr>
<tr>
<td>+3</td>
<td>+4</td>
<td>+5</td>
<td>+6</td>
<td>+7</td>
<td>+8</td>
</tr>
<tr>
<td>+9</td>
<td>+10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Instructions to Subjects. I am going to read you some instructions. I am not permitted to say anything which is not in the instructions nor can I answer any questions about this experiment. OK?

We are in the process of developing a test of empathy. This test is designed to show how well a person is able to put himself into someone else's place. I will show you a series of photographs. For each one I want you to judge whether the person pictured has been experiencing success or failure. To help you make more exact judgements you are to use this rating scale. As you can see the scale runs from -10 to +10. A rating of -10 means that you judge the person to have experienced extreme failure. A rating of +10 means that you judge the person to have experienced extreme success. A rating of -1 means that you judge the person to have experienced mild failure, while a rating of +1 means that you judge the person to have experienced mild success. You are to rate each photo as accurately as you can. Just tell me the rating you assign to each photo. All ready? Here is the first photo. (No further explanation may be given, although all or
part of the instructions may be repeated.)

From the original 57 photos, 10 were selected for presentation to male subjects and 10 were selected for presentation to female subjects. All 20 photos were rated on the average as neither successful nor unsuccessful and for each the mean rating evoked fell between +1 and -1. The distributions of ratings evoked by each of the photos were also symmetrical (Rosenthal, 1966, p. 143 f.).
APPENDIX B

INSTRUCTIONS TO SUBJECTS IN THE VERBAL CORRECT FEEDBACK CONDITION

We are in a process of developing a test of empathy. This test is designed to show how well a person is able to put himself in someone else's place. The Experimenter will show you a series of photographs. For each one I want you to judge whether the person pictured has been experiencing success or failure. You are to rate each person as accurately as you can.

In the experimental room you will sit opposite the Experimenter at a table. In front of you will be a panel. On the panel is a scale that looks like this.

<table>
<thead>
<tr>
<th>FAILURE</th>
<th>SUCCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

To help you make more exact judgements you are to use this rating scale. As you can see the scale runs from 1 to 8. A rating of 1 means that you judge the person to have experienced extreme FAILURE. A rating of 8 means that you judge the person pictured to have experienced extreme SUCCESS. A rating of 4 means that you judge the person to have experienced mild FAILURE while a rating of 5 means that you judge the person to have experienced mild SUCCESS.

To indicate your response just tell the Experimenter the number that corresponds to your rating for each photo that is presented. Give him only the number of your rating.
APPENDIX C

INSTRUCTIONS TO SUBJECTS IN THE VERBAL REVERSED FEEDBACK CONDITION

We are in the process of developing a test of empathy. This test is designed to show how well a person is able to put himself in someone else's place. The Experimenter will show you a series of photographs. For each one I want you to judge whether the person pictured has been experiencing success or failure. You are to rate each person as accurately as you can.

In the experimental room you will sit opposite the Experimenter at a table. In front of you will be a panel. On the panel is a scale that looks like this.

SUCCESS

1 2 3 4 5

FAILURE

6 7 8

To help you make more exact judgements you are to use this rating scale. As you can see this scale runs from 1 to 8. A rating of 1 means that you judge the person to have experienced extreme SUCCESS. A rating of 8 means that you judge the person to have experienced extreme FAILURE. A rating of 4 means that you judge the person to have experienced mild SUCCESS while a rating of 5 means that you judge the person to have experienced mild FAILURE.

To indicate your response just tell the Experimenter that number that corresponds to your rating for each photo that is presented. Give him only the number of your rating.
APPENDIX D

INSTRUCTIONS TO SUBJECTS IN THE VERBAL NO FEEDBACK CONDITION

We are in the process of developing a test of empathy. This test is designed to show how well a person is able to put himself in someone else's place. The Experimenter will show you a series of photographs. For each one I want you to judge whether the person pictured has been experiencing success or failure. You are to rate each photo as accurately as you can.

In the experimental room you will sit opposite the Experimenter at a table. In front of you will be a panel. On the panel is a rating scale that looks like this.

<table>
<thead>
<tr>
<th>FAILURE</th>
<th>SUCCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

To help you make more accurate judgements you are to use this scale. As you can see the scale runs from 1 to 8. A rating of 1 means that you judge the person to have experienced extreme FAILURE. A rating of 8 means that you judge the person to have experienced extreme SUCCESS. A rating of 4 means that you judge the person to have experienced mild FAILURE while a rating of 5 means that you judge the person to have experienced mild SUCCESS.

To indicate your responses press the button on your panel under the number that corresponds to your rating. A light on your panel will go on to indicate that your response has been recorded. Do not tell the Experimenter your rating.
APPENDIX E

INSTRUCTIONS TO SUBJECTS IN THE NONVERBAL CORRECT FEEDBACK CONDITION

We are in a process of developing a test of empathy. This test is designed to show how well a person is able to put himself in someone else's place. The Experimenter will show you a series of photographs. For each one I want you to judge whether the person pictured has been experiencing success or failure. You are to rate each photo as accurately as you can.

In the experimental room you will sit opposite the Experimenter at a table. In front of you will be a panel. On the panel is a rating scale that looks like this.

<table>
<thead>
<tr>
<th>FAILURE</th>
<th>SUCCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

To help you make more accurate judgements you are to use this scale. As you can see the scale runs from 1 to 8. A rating of 1 means that you judge the person to have experienced extreme FAILURE. A rating of 8 means that you judge the person to have experienced extreme SUCCESS. A rating of 4 means that you judge the person to have experienced mild FAILURE while a rating of 5 means that you judge the person to have experienced mild SUCCESS.

To indicate your choice press the button on your panel under the number that corresponds to your rating. A light on your panel will go on to indicate that your response has been recorded.

When you enter the test situation, please do not talk to the Experimenter either before or during the presentation of the photographs. Sit down in the chair opposite the Experimenter and he will present the first photograph. Save any questions you may have until after the experiment.
APPENDIX F

INSTRUCTIONS TO SUBJECTS IN THE NONVERBAL REVERSED FEEDBACK CONDITION

We are in a process of developing a test of empathy. This test is designed to show how well a person is able to put himself in someone else's place. The Experimenter will show you a series of photographs. For each one I want you to judge whether the person pictured has been experiencing success or failure. You are to rate each photo as accurately as you can.

In the experimental room you will sit opposite the Experimenter at a table. In front of you will be a panel. On the panel is a rating scale that looks like this.

SUCCESS

1 2 3 4 5 6 7 8

FAILURE

To help you make more accurate judgements you are to use this scale. As you can see the scale runs from 1 to 8. A rating of 1 means that you judge the person to have experienced extreme SUCCESS. A rating of 8 means that you have judged the person to have experienced extreme FAILURE. A rating of 4 means that you judge the person to have experienced mild SUCCESS while a rating of 5 means that you judge the person to have experienced mild FAILURE.

To indicate your choice press the button on your panel under the number that corresponds to your rating. A light on your panel will go on to indicate that your response has been recorded.

When you enter the test situation, please do not talk to the Experimenter either before or during the presentation of the photographs. Sit down in the chair opposite the Experimenter and he will present the first photograph. Save any questions you may have until after the experiment.
APPENDIX G

INSTRUCTIONS TO SUBJECTS IN THE NONVERBAL NO FEEDBACK CONDITION

We are in a process of developing a test of empathy. This test is designed to show how well a person is able to put himself in someone else's place. The Experimenter will show you a series of photographs. For each one I want you to judge whether the person pictured has been experiencing success or failure. You are to rate each photo as accurately as you can.

In the experimental room you will sit opposite the Experimenter at a table. In front of you will be a panel. On the panel is a rating scale that looks like this.

<table>
<thead>
<tr>
<th>FAILURE</th>
<th>SUCCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

To help you make more accurate judgements you are to use this scale. As you can see the scale runs from 1 to 8. A rating of 1 means that you judge the person pictured to have experienced extreme FAILURE. A rating of 8 means that you judge the person to have experienced extreme SUCCESS. A rating of 4 means that you judge the person to have experienced mild FAILURE while a rating of 5 means that you judge the person to have experienced mild SUCCESS.

To indicate your response press the button on your panel under the number that corresponds to your rating. A light on your panel will go on to indicate that your response has been recorded.

When you enter the test situation, please do not talk to the Experimenter either before or during the presentation of the photographs. Sit down opposite the Experimenter and he will present the first photograph. Save any questions you may have until after the experiment.