THE UNIVERSITY OF MANITOBA

EXPERIMENTER EXPECTANCY EFFECT AS A FUNCTION OF SUBJECTS' VOLUNTEER STATUS AND EVALUATION APPREHENSION

by

HAMIDA AKHTAR BEGUM

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

WINNIPEG, MANITOBA

February 1973

EXPERIMENTER EXPECTANCY EFFECT AS A FUNCTION OF SUBJECTS' VOLUNTEER STATUS AND EVALUATION APPREHENSION

BY: Hamida Akhtar Begum

A dissertation submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

ABSTRACT

The study examined the effects of subject's volunteer status and evaluation apprehension on experimenter expectancy effect. Two groups of subjects, volunteers (16 males, 32 females) and coerced subjects (39 males, 41 females), were recruited from different sections of the Introductory Psychology course at the University of Manitoba. One half of the subjects from each group was induced with high and the other half with low evaluation apprehension by means of written descriptions about the study. The experimental task consisted of ten chromatic colors which were rated in terms of color-richness on a 20-point scale. Subjects were tested on this task by one of the 18 male experimenters who were previously biased to expect either high (+5) or low (-5) average ratings from their subjects. In addition, to measure subjects' need for approval and attitude toward psychological research, their responses to the Marlowe-Crowne Social Desirability Scale and the Psychological Research Survey were obtained before and after the experimental task, respectively.

As predicted, volunteers scored higher on both measures showing significantly more positive attitudes toward research and, although not significant, stronger need for approval than coerced subjects. Assuming that these differences in attitudes and needs would have differential effects on their response to the color-judgment task, it was predicted that experimenter expectancy effect, i.e., the difference between the ratings obtained by experimenters under high and low expectancies, would be greater for volunteers than for coerced subjects. It was also predicted that the expectancy effects for volunteers would be equally greater under both high and low evaluation apprehension; whereas, with coerced subjects

it would be greater under high evaluation apprehension than under low evaluation apprehension. Finally, a significant experimenter expectancy effect was predicted across all treatment conditions.

The results of an analysis of variance of subjects' mean ratings did not support these hypotheses. Although the mean ratings were in the predicted direction, volunteers did not show significantly greater expectancy effect than coerced subjects. Similarly, the expectancy effect was not found to vary significantly with subjects' volunteer status and evaluation apprehension. It was found, however, that volunteers' responses were enhanced by heightened evaluation apprehension more than those of coerced subjects. Finally, although the overall mean ratings of subjects were in the direction of experimenter expectancies, the difference between the ratings obtained in high and low expectancy conditions was not significant.

However, a post-hoc analysis of these data according to the subjects' sex revealed a significant interaction of experimenter expectancy, volunteer status and sex of subjects. The expectancy effect was greater with female volunteers and male coerced subjects than with female coerced subjects and male volunteers. In addition, a significant interaction between sex of subjects and evaluation apprehension was obtained which indicated that the ratings of female subjects were enhanced under high evaluation apprehension more than those of male subjects. Further post-hoc analyses of these data according to subjects' level of attitude toward psychological research and need for approval failed to show any additional significant effects.

Thus, although the predictions were not supported, the results of the post-hoc analysis according to sex of subjects indicated that subjects' sex and volunteer status may combine in a complex way to determine

their responses to experimenter expectancy and evaluation apprehension. Further research was suggested to test these relationships.

ACKNOWLEDGEMENTS

It gives the author a great pleasure to express her deepest gratitude to Dr. J. G. Adair for his constant guidance and encouragement in accomplishing this research. Sincerest thanks are also extended to Dr. D. Perlman who has helped the author in many commendable ways. The author also wishes to express her appreciation to Dr. R. Altemeyer for his valuable suggestions and criticisms of the manuscript.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

LIST OF TABLES

LIST OF TABLES (CONTINUED)

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Recent research on the social psychology of the psychological experiment has focused attention on the methodology of the behavioral sciences. This research has shown that the characteristics and expectations of the experimenter as well as of the subjects (Rosenthal, 1966; 1969) and the unintended cues present in the experimental situation (Orne, 1962; 1969) may influence the validity of empirical data. In other words, it has indicated that the experimental situation is often not what behavioral scientists intended to create. Instead, it has been suggested, "the process of collecting data about human behavior is itself a social process and shares features in common with other situations in events of human interaction (Riecken, 1962; p. 26)".

Rosenthal and his colleagues have presented evidence that the expectancy or hypothesis held by an experimenter may influence the interaction between the subject and the experimenter in such a way that the hypothesis is confirmed (Rosenthal, 1966; 1969). The typical paradigm of their experiments had experimenters administer a person perception task to subjects who were asked to rate ten photographed faces on a scale of 20-points ranging from -10 (extreme failure) to +10 (extreme success). Bias was introduced by inducing experimenters to expect certain average ratings of the photographs. One group of experimenters received a positive (+5) expectancy for the ratings while the remaining group received a negative (-5) expectancy. The effects of experimenter expectancy was determined from a significant difference between mean ratings obtained from subjects tested by positively

and negatively biased experimenters.

Within this basic paradigm, Rosenthal and his colleagues have demonstrated the effect of the experimenter expectancy with human subjects (Rosenthal & Fode, 1963b; Fode, 1965; Rosenthal & Laszlo, 1967), as well as with laboratory animals (Rosenthal & Lawson, 1964; Rosenthal & Fode, 1963a). In addition, the experimenter expectancy effect has been demonstrated in a variety of tasks and situations such as motor performance (Johnson, 1970a, 1970b), learning (Kennedy, Cook, & Brewer, 1970), clinical testing (Marwit & Marcia, 1967; Masling, 1965; 1966), sensory deprivation research (Raffetto, 1968) and even in classroom performance of students (Rosenthal & Jacobson, 1968; Beez, 1968).

In spite of this extensive empirical support, the generality of experimenter expectancy effect has been questioned (Barber & Silver, 1968). Some investigators have been unable to demonstrate the effect (Pflugrath, 1962; Wartenberg-Ekran, 1962; Friedman, 1967). Referring to these and many other studies, Barber and Silver (1968) suggest:

> "Further research is needed to determine under what circumstances we are most likely to obtain an experimenter bias effect. The parameters to be considered include the characteristics of the participating individuals (principal investigator, experimenters and subjects), the methods used to induce expectancy biases and the nature of the experimental task (p. 26)."

The present study was designed to investigate some of the processes involved in this phenomenon by focusing on the "participating individuals", particularly the subjects.

Mediation of Experimenter Expectancy Effect: Subject's Evaluation Apprehension

The most parsimonious explanation of the phenomenon is that the experimenter's expectancies lead him to act differently toward subjects, and that this subtle, unintentional, but discriminatory influence elicits expectancy

supporting behaviors. This is in line with Orne's (1962) contention that numerous stimuli in the experimental situation suggest to the subject how the experimenter would like him to respond. He labelled these "cues which convey an experimental hypothesis to the subject (Orne, 1962, p. 799)" as the demand characteristics of the situation. Within this framework, experimenter expectancy may function through systematically different 'behavioral demand characteristics' presented to subjects by their experimenter.

While Rosenthal is particularly concerned with the experimenter's side of the experimenter-subject interaction, Orne (1962), Riecken (1962) and Rosenberg (1965) have focused on the subject as a comparatively active, yet obedient recipient of 'implicit cues' from the experimenter. For example, Riecken (1962) suggested that subjects of psychological research actively strive to discern the purpose of the experiment so that they may maximize their rewards by "putting their best foot forward". Rosenberg (1965; 1969) elaborated this view and stressed the notion of "subject as seeker"-- a seeker of experimenter's positive evaluation which enables him to maintain or enhance his self-esteem. He proposed that experimental situations, in varying degrees, arouse within the subject an "anxiety-toned concern that he win a positive evaluation from the experimenter (Rosenberg, 1965; p. 29)", a process he designated as "evaluation apprehension". In his view, such a process may have a systematic effect on research data.

This proposition was examined in a series of experiments in which the subjects' apprehension was aroused by written communications prior to their performance. The first of these experiments (Rosenberg & Mulry; reported in Rosenberg, 1969) used a person perception task that required subjects to rate on a 20-point scale the degree of their liking for some photographed

persons. Subjects were made to feel apprehensive or ego-involved in their performance by informing them that psychologically mature and healthy persons show greater liking for strangers than do immature people. To another group of subjects the opposite was communicated, i.e., psychologically immature and comparatively unhealthy persons show greater liking for strangers. A control group of subjects received a neutral communication. The results showed that the ratings obtained from the subjects in the 'liking' treatment were significantly higher than those in the 'dislike' treatment. The ratings also differed significantly from those of the control group in the predicted order, i.e., attraction scores were highest in the 'liking' treatment, lowest in the 'disliking' treatment, and intermediate in the control condition.

Within the same experimental paradigm, another study (Rosenberg, 1969) was undertaken with problem solving tasks. Evaluation apprehension was aroused in two experimental groups by telling subjects either that psychologically healthy and mature persons would be less efficient and would find less pleasure than the psychologically immature persons in performing the tasks, or the exact opposite. A control group of subjects received no information. The results showed that the three groups differed from one another in the predicted direction. The subjects who were to believe that mature people tend to be inefficient completed significantly fewer tasks and solved significantly fewer problems than did the opposite experimental group.

In another study (Rosenberg, Corsi, & Holmes; reported in Rosenberg, 1969) subjects were required to tap a key with index fingers of both hands. Evaluation apprehension was aroused by telling subjects that people with higher intelligence would more likely perform better with both index

fingers than less intelligent people, i.e., it was implied that the difference between the number of tappings with right and left fingers would be less with people of high intelligence. Apprehension was further strengthened by prior testing of the subjects on three ability tests involving verbal and symbolic skills. Control subjects received neither apprehensive information nor ability tests. As predicted, the mean difference between the number of taps with the right and left index finger was significantly lower in the evaluation apprehension group than in the control group.

These studies established that subject's responses are determined by his perception of how his responses will be judged and his concern over winning a positive evaluation from the experimenter. Referring this to the experimenter expectancy research, Rosenberg hypothesized that,

> "evaluation apprehension may well be involved in experimenter expectancy effect, i.e., a state of concern over whether the experimenter will judge him as 'normal' or 'abnormal' may affect the way in which the subject perceives the experimental situation..... The subject who is possessed of a concern over evaluation apprehension may well be more closely and accurately attuned to such indirect communication or he may be more motivated to act upon the basis of what has been indirectly communicated (Rosenberg, 1969, p. 322)."

This hypothesis was tested in a series of studies that manipulated levels of subjects' evaluation apprehension in a typical experimenter expectancy paradigm. Using Rosenthal's person perception task, Minor (1970) demonstrated a significant interaction between evaluation apprehension and experimenter expectancy conditions. He found that the experimenter's expectancies led to hypothesis confirming responses only when subjects were made personally concerned with their performance.

Similar results were obtained by Duncan, Rosenberg and Finkelstein (1969) in another study employing the person perception task. The design

was similar to Minor's (1970) except that in this study the subjects with either high or low evaluation apprehension were exposed to the taperecorded instructions that were "slightly shaded" in either a positive (i.e., "success" stressing) or a negative (i.e., "failure" stressing) direction. In addition, this study included a control group for evaluation apprehension. As predicted, the results showed a strong interaction between differential non-verbal cueing (experimenter expectancy) and evaluation apprehension. The paralinguistic cues influenced subject's responses such that the extent of the influence was in direct proportion to the strength of evaluation apprehension.

Although these two studies supported Rosenberg's prediction, Blake and Heslin (1971) did not find evaluation apprehension to be that important in demonstrating experimenter expectancy effect in a different kind of picture-rating task. They found that subjects' data varied with the experimenter's expectancy but that this effect was no greater under high evaluation apprehension than under a non-evaluative set. Although subjects gave more socially desirable responses under high evaluation apprehension, even this effect diminished in the later trials of the experimental session. In comparison with the earlier studies showing the effect of evaluation apprehension on experimenter expectancy effect, it should be noted that the Blake and Heslin procedure was quite different. For example, the presence of confederates who confirmed the experimenter's hypothesis and the attempted manipulation of low evaluation apprehension by leading subjects to believe the task was a measure of contemporary consumers' appraisal of the market value of the pictures, may have sufficiently altered the experiment to find the effect of evaluation apprehension. On the other hand, since they found a significant experimenter expectancy effect, their experimenters may have

presented their expectancies so clearly that any effect of evaluation apprehension were unnecessary to achieve the experimenter expectancy effect.

Recently, Johnson (1970b)attempted to replicate Minor's findings by using Stevenson's marble-dropping task. Experimenter expectancies were induced at two levels by indicating to the experimenters which subjects would show better performance as they were alleged to be brighter than others. Evaluation apprehension was induced in subjects at high, medium, and low levels by giving them an intelligence test and then providing an information sheet intended to induce evaluation apprehension prior to the experimental task; giving the intelligence test only; and providing an information sheet intended to induce no evaluation apprehension, respectively. In addition, experimenters' concerns over the experiment were varied at three levels by employing three conditions of expectancy inducements. The results confirmed Rosenberg's hypothesis only when experimenter's concern over the outcome of the experiment was high. Under low experimenter concern, no significant interaction between expectancy and evaluation apprehension was obtained. It was concluded that a minimum of both experimenter outcome conern and subject performance concern must be present for expectancy mediation. However, another reason for the failure to replicate Minor's findings cited by Johnson is that evaluation apprehension may not have been effectively induced at differential levels.

Thus, it appears that while subjects' evaluation apprehension has been found to be an important condition for experimenter expectancy effect, the process involved in the mediation of this effect still remains vague. Apart from experimentally induced concern, i.e., evaluation apprehension, none of the aforementioned studies gave special consideration to the subject's

motivational background. Certainly it should be considered that when a subject steps into an experimental room he has a certain degree of motivation to help scientific research and is influenced by various internal and external factors, such as, his predisposing personality characteristics, personal interest and attitudes toward psychological research, and the situational factors that led him to participate in the experiment. This motivational background conceivably determines the way he would respond and his sensitivity to any experimental manipulation, including experimenter expectancy and evaluation apprehension. Recently, Adair (1972) has pointed to the importance of the subject's predisposing characteristics. He contends that the subject's pre-experiment attitudes predispose him to respond positively or negatively toward experimental manipulations. Thus, a thorough understanding of experimenter expectancy phenomenon requires careful consideration of subject selection and motivation variables and their interaction with evaluation apprehension and expectancy effects.

Volunteer Status of Subjects

An important source ofsubject'smotivational differences in behavioral research is the manner in which they are selected. While introductory psychology students are often coerced into participation, in many studies volunteers are solicited. A recent survey of psychological experiments (Jung, 1969) shows that 54% of all subjects were recruited through some form of voluntary basis. This survey also shows that 80% of all research obtained subjects from introductory psychology courses. Referring to these facts Rosenthal and Rosnow (1969) have remarked, "The existing science of human behavior may be largely the science of these sophomores who both (a) enroll in psychology courses and (b) volunteer to participate in behavioral research (pp. 59-60)".

The problem with the use of volunteers is primarily motivational. Those who are eager to participate have been found to differ from non-volunteers on a number of psychological dimensions. For example, in one study (McDavid, 1965) subjects recruited under four degrees of external inducement were tested on an approval-seeking motivation scale. It was found that the 'captive' subjects who were tested without forewarning (in a regular classroom) received the lowest and the most widely dispersed scores; those who participated as an academic requirement, i.e., 'required' subjects, and 'volunteers' who received extra grade credits for participation showed intermediate scores; and subjects who volunteered with minimal external inducement obtained the highest mean and the least dispersed scores. The groups differed significantly as a function of the conditions of recruitment, the more 'voluntarism' the higher their approval need.

In general, research (Rosnow & Rosenthal, 1970) has shown that volunteers and non-volunteers are distinguishable on a number of characteristics. Volunteers have been found to have greater intellectual ability, interest and motivation than non-volunteers. Furthermore, volunteers tend to be less authoritarian, less maladjusted, and, as McDavid found, have a greater need for social approval. With lesser confidence it has also been found that volunteers tend to be more sociable, arousal seeking, unconventional and first born children. These conclusions, based on numerous studies, indicate that volunteers can never be a random sample of the entire population. It should be noted that the reason the volunteer subjects are not representative of the larger population is not simply because they have come forward to participate in an experiment, rather it is because they differ from non-volunteers in their perception of and reaction to the cues

that impart information about the purpose of the study. Thus, subject selection poses a problem when recruitment procedures lead to samples of subjects who possess different characteristics that are relevant to the response to the information cues of the experiment.

Although attempts have been made to minimize the problem of volunteer bias by constituting subject pools through which undergraduate students in introductory psychology are required to participate in a certain number of experiments, the problems still remain. Since the students are free to decide when they will fulfill their requirement, the problem of volunteer bias emerges from their decision to participate either earlier to later in the term. For example, it has been found (Adair, 1970) that subjects participating in experiments early in the year (presumably volunteers) generally had more positive attitudes toward psychological research than those who participated later (presumably non-volunteers). Thus, the problem of volunteer bias is present in some form in most research with human subjects.

The implications of volunteer bias for non-representativeness are considerable. The problem is most significant when the personality and motivational differences interact with experimental manipulations. For example, Rosenthal and Rosnow (1969) noted that "volunteer status may interact with experimental variables in such a way as to increase the probability of inferential errors of the first and the second kind.[1] (p. 112)." They demonstrated this in a series of attitude change studies comparing the behavior of both volunteers and non-volunteers. In a traditional opinion change experiment (Rosnow and Rosenthal, 1966), it was found that the volunteers consistently responded in accord with the

---

[1] Type 1 error refers to the rejection of a true null hypothesis, Type II error refers to the acceptance of a false null hypothesis.

experimental hypothesis. When the communication was positive, volunteers became more positive compared with a control group and with non-volunteers; and when the communication was negative, volunteers became more negative. Thus, volunteers were more strongly motivated and cooperative than non-volunteers.

Volunteers have also been found to be more sensitive to the demand characteristics of the experiment than non-volunteers. A study (Rosnow, Rosenthal, McConochie, & Arms, 1969) showed that when the communication was too obtrusive and obvious, volunteers did not cooperate with the persuasive attempt and their opinion did not differ from those of non-volunteers. In contrast, when the communication was subtle, volunteers confirmed the experimental hypothesis significantly more than non-volunteers. In a more recent study (Rosnow & Suls, 1970) a significant interaction between volunteer status and attitude pre-testing was obtained that indicated the facilitative effects of pre-testing for volunteers and depressive effects for non-volunteers. Further evidence of volunteer bias comes from Horowitz (1969) who found that volunteers complied to a significantly greater extent with the recommendations of a persuasive message and were affected to a significantly greater degree by the arousal manipulations than non-volunteers.

Another area in which volunteer effects have been explored to some extent is verbal conditioning. Goldstein, Rosnow, Goldstadt and Suls (1970) found that volunteers showed a significantly greater increase in critical (reinforced) responses than non-volunteers. Recently, Cox and Sipprelle (1971) demonstrated operant verbal conditioning of acceleration and deceleration of heart rate occurred significantly more often among

volunteers but not with participants who were obtained from a compulsory subject pool. In another study (Black, Schumpet, & Welch, 1972) involving pursuit rotor performance, volunteers were found to have a greater accuracy and perseverence than the subjects who were coerced to participate as a course requirement. Considering this research, it appears that volunteers are more sensitive to the demand characteristics operating in a given situation and are often more cooperative with experimenters than non-volunteers.

Volunteer Problem in Experimenter Expectancy and Evaluation Apprehension Research

As in other research, experiments on experimenter expectancy and evaluation apprehension also involve the problem of volunteer bias. In experimenter expectancy research, "most of the subjects were volunteers" and the "greatest single contributing course was introductory psychology (Rosenthal, 1966, p.308)". The experiments involving evaluation apprehension also employed volunteer subjects. For example, Minor (1970) reported using college student volunteers. Duncan, Rosenberg, and Finkelstein (1969) indicated that their subjects volunteered following telephone requests for participation. Rosenberg (1969) also reported that his subjects were volunteers.

It should be noted, however, that the precise definition of the word 'volunteer' is often vague in this research. 'Volunteer subject' means different things to different investigators. In addition, journal articles hardly provide a detailed description of subject selection procedures. Often the description is confined to a phrase 'the subjects volunteered' and in many cases nothing is mentioned about subject recruitment procedures. As a result, it is difficult to isolate the effects of this variable in the

research on experimenter expectancy and evaluation apprehension. However, research in which volunteering was manipulated as an independent variable (Rosenthal & Rosnow, 1969) suggests that experimenter expectancy effects could be differentially affected by subjects' volunteer status.

Experimenter expectancy research is essentially a social influence situation where certain cues, transmitted by the experimenters with different expectancies, are perceived and interpreted by the subjects. Since volunteers have been found to have more positive attitudes toward psychological research (Adair, 1970) and greater sensitivity to the demand characteristics than non-volunteers (Rosnow, et al., 1969; Rosnow & Suls, 1970) they are likely to be more efficient in receiving these cues. In addition, their greater need for social approval relative to that of non-volunteers (McDavid, 1965; Rosnow & Rosenthal, 1970) predisposes them to cooperate with what the experimenter wants them to do. Thus it could be expected that experimenter expectancy effect would be greater when subjects volunteer than when they are coerced to participate.

The effects of volunteer status on the results of expectancy research also seem to be dependent upon the levels of evaluation apprehension in the study. Being predisposed to be sensitive to the experimental cues and approval seeking, volunteers are likely to confirm experimenter's hypothesis more than non-volunteers under low evaluation apprehension. With the arousal of high evaluation apprehension, however, non-volunteers would become attuned to experimenter's cues and perhaps emit more responses confirming the expectancy. The responses of volunteers, on the other hand, are less likely to be affected by high evaluation apprehension since they are already sensitive and cooperative. In other words, there may not be a

significant difference in their responses under high and low evaluation apprehension conditions. Compared to the less positively motivated non-volunteers, however, volunteers would confirm the experimenter's hypothesis more strongly under heightened evaluation apprehension as well. From this view, then, an interaction between experimenter expectancy, volunteer status and evaluation apprehension would be expected. Thus, the present study was designed to examine the effects of volunteer status and evaluation apprehension on experimenter expectancy effect.

Statement of the Problem

The results of experimenter expectancy research are obscured by the use of volunteer and non-volunteer subjects who are solicited in various ways. The problem arises from the fact that persons who volunteer for an experiment often differ from those who are coerced to participate on important psychological characteristics which make them differentially sensitive to the experimental cues and cooperative with the experimenter (Rosenthal & Rosnow, 1969). The present study was, therefore, designed to examine the effects of subject's volunteer status on experimenter expectancy research under conditions of manipulated evaluation apprehension.

Assuming that volunteers are more strongly motivated toward research than non-volunteers, it was expected that they would be more sensitive to the unintended cues transmitted by experimenters with specific expectancies and would be more likely to confirm the hypotheses. In other words, volunteers were expected to show greater expectancy effect than non-volunteers. Under the inducement of heightened evaluation apprehensiion, however, volunteers are less likely to be further aroused since they are already predisposed to be sensitive and cooperative. Thus, with volunteer subjects,

it is unlikely that there would be any significant difference in expectancy effects between high and low levels of evaluation apprehension. In contrast, non-volunteers would, perhaps, emit more expectancy confirming responses under high evaluation apprehension. They would be motivated to guess what is expected of them to maintain or enhance their self-esteem and this would guide their responses. In other words, experimenter expectancy effect would differ as a function of subjects' volunteer status and evaluation apprehension.

To test the validity of this contention, the following experiment was designed. Two groups of subjects, 'volunteers' and 'coerced' subjects, were aroused with either high or low evaluation apprehension, and were tested on a color-judgment task (Shames, 1971) under the supervision of an experimenter with either a high (+5) or low (-5) expectancy about the subjects' average ratings. The judgments of colors were made on a 20-point scale of color richness.

This task was selected for several reasons. First, a previous study of experimenter expectancy effect has successfully obtained the phenomenon with this task (Shames, 1971). In view of the elusive nature of experimenter expectancy effect of Rosenthal (1969) and the objectives of this experiment, this was an important criterion in the selection of the task. In addition, this task was similar to the person perception task often used in experimenter expectancy research with evaluation apprehension manipulation (Minor, 1970; Rosenberg et al., 1969).

The assumption that volunteers differ from coerced subjects in motivational strength and compliant tendency was tested by assessing their attitude toward psychological research and need for social approval. These

were measured by the Psychology Research Survey (Adair & Fenton, 1970) and Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1964), respectively.

With regard to the major variables, it was hypothesized that:

1. There will be a significant interaction between experimenter expectancy and subjects' volunteer status, i.e., experimenter expectancy effect, i.e., a significant difference between mean color ratings obtained under high and low expectancy conditions, will be greater with volunteers than with coerced subjects.

2. There will be a significant triple interaction among experimenter expectancy, volunteer status and evaluation apprehension such that experimenter expectancy effect for coerced subjects will be greater under high evaluation apprehension than under low evaluation apprehension, while for volunteer subjects this difference will not be obtained. Volunteers will consistently show a greater expectancy effect than that of coerced subjects under both high and low evaluation apprehension conditions.

3. There will be a significant experimenter expectancy effect, i.e., subjects under high expectancy will show higher ratings of colors than subjects under low expectancy across all treatment conditions.

CHAPTER II

METHOD

Subjects

The subjects were 55 male and 73 female students enrolled in the Introductory Psychology course at the University of Manitoba. As a course requirement, every student was expected to participate in psychological experiments to obtain five hours of credit for which he received 5% of his final grade in the course. The subjects were obtained from this population by two different sampling methods, used to procure subjects who differed in their perceived freedom to participate in the study. Thus one set of subjects were obtained from the compulsory subject pool in the usual fashion, i.e., they were 'coerced subjects', while the others were 'volunteers'.

Coerced subjects. The coerced subjects were 43 male and 46 female students who signed for the experiment to obtain one hour of credit. Of these four males and five females did not appear, leaving a final sample of 39 male and 41 female subjects. The principal investigator obtained these subjects by going to four randomly selected sections of the Introductory Psychology course (each section enrolling 70 to 80 students) and displayed the sign-up booklet to the students without making any special appeal for participation.

Volunteers. The volunteers were 23 male and 35 female students who were solicited by the principal investigator through a request to serve as subjects on a completely voluntary basis. Of these, seven males and three

females did not appear, leaving a final sample of 16 male and 32 female subjects. To obtain these subjects, a special request (Appendix A) was made in another fifteen sections of the same course. As with the coerced subjects the principal investigator personally appeared in the classroom and told the students that she needed volunteers to help complete an experiment for her thesis. She added that it was not possible for her to offer any experimental credit as she had already tested the subjects assigned to her. Therefore, the subjects who agreed to help her would have done so on a truly voluntary basis.

Experimenters

From a larger number of volunteers, 18 male students from advanced courses in Statistics and Engineering at the University of Manitoba, were selected to serve as experimenters. Each one was contacted by telephone and was offered $3.00 for participation. They were requested to assist in a psychological study for the principal investigator's Doctoral thesis. Potential experimenters were told that their recruitment was necessary because the appearance and manner of an experimenter from a different culture, such as the case with the principal investigator, could affect the data. Furthermore, it was indicated that assistance could not be solicited from students in the Psychology Department as most of them were busy with their own experiments. None of the experimenters appeared to be suspicious of this rationale. Those who agreed to participate were asked to come fifteen minutes before the experimental session so that the task and details of the study could be explained.

Design

The study consisted of a 2 x 2 x 2 factorial design representing two

levels of experimenter expectancy (high and low), two levels of evaluation apprehension (high and low), and two levels of volunteer status of subjects (volunteers and coerced). Subjects were randomly assigned to each cell. Since experimenters were to test subjects in all conditions, an attempt was made to randomly assign an approximately equal number of volunteers and coerced subjects under each experimental treatment. However, such equality could not be maintained as some subjects did not appear. The number of subjects tested by each experimenter ranged from three to nine, with most experimenters testing at least six subjects.

Task

The experimental task, devised by Shames (1971), consisted of a set of ten chromatic colors which were to be rated on a 20-point scale of color richness (Appendix B) ranging from -10 (extremely color poor) to +10 (extremely color rich). The scale was similar to that used in Rosenthal's (1966) person perception task. The color stimuli were ten 2 1/2 x 1 3/4 inch arborite chips developed by Canadian Industrial Limited (C.I.L.) which were mounted separately on ten 3 x 5 inch white cards.

Training of the Experimenters

As requested, each experimenter arrived approximately fifteen minutes before the arrival of the first subject. He was greeted by the principal investigator and was taken to one of the three research rooms used for the experiment. Each experimenter was first asked to silently read the instructions to the color-judgment task (Appendix C) so that he could understand what the subjects were supposed to do. The instructions included information about how to use the rating scale in judging colors as to their richness or poorness. In order to make it clear, the principal investigator demonstrated the cards with chromatic colors, the rating scale and the

procedure to be followed in testing each subject. After greeting the subject, the experimenter was to ask and record his or her name on the recording sheet (Appendix D) and then read the instructions clearly and distinctly. He was warned to say nothing about the study other than what was printed in the instructions. However, if the subjects did not understand or asked a question, he was allowed to repeat the appropriate section of instructions. Following this, the experimenter was asked to present each card according to its serial number (1 to 10) and record the ratings assigned by the subject. At the conclusion of the color-judgment task he was to send each subject to the principal investigator in another room. He was told that there the subjects would answer some post-experimental questionnaires and then receive information about the purpose of the experiment from the principal investigator.

Inducement of Experimenter Expectancy

In the course of explaining the task, the experimenters were induced to expect results from different subjects (Appendix E). This was achieved by leading them to believe that the color-judgment task was a way of testing the subjects' emotional characteristics, i.e., people with different emotional characteristics rate the colors differently. Referring to some non-existent data about the emotional characteristics of the subjects participating in this study, it was indicated that some subjects were expected to rate the colors positively (+5) while others would rate them negatively (-5). The names of the persons for whom either positive or negative ratings were expected were also indicated on a list of subjects to be tested provided for each experimenter.

Procedure

The experiment was conducted in eighteen separate sessions in the last two weeks of March and the first week of April, 1972. Each session took place in three experimental rooms situated at separate locations on the fifth floor of the Psychology Department of the University of Manitoba.

As each subject arrived, he/she was greeted by an assistant of the principal investigator in Room I. This assistant was one of four persons, two Graduate and two Fourth Year Honours students in Psychology, who had volunteered to help the principal investigator as their schedules permitted. Each was kept blind to the purpose of the experiment. The assistant told the subject on arrival that the experimenter was busy with another subject and asked if he/she would complete a brief questionnaire. Since no subject refused, the Marlowe-Crowne Social Desirability (S-D) scale was then administered.

Following this, one-half of the subjects were asked to read a printed description intended to induce high evaluation apprehension (Appendix F). It described the color-judgment task as a test of emotional adjustment. The readers were led to believe that much previous research had shown that this test could determine emotionally maladjusted people from a college population. The experiment they were participating in was described as an attempt to replicate those findings.

The other half of the subjects read another description intended to induce low evaluation apprehension (Appendix G). In this case, the experiment was described as part of the study on factors related to color-perception. Subjects were led to believe that they belonged to the control group

and their performance on the subsequently presented color-judgment task would provide a standard for comparison of ratings for those in the experimental groups. These description sheets were assigned to subjects according to a list of names provided for the assistants. In this way the number of subjects in each of the experimental group could be controlled.

After reading the description, the subject was led to Room II where an experimenter administered the color-judgment task. The experimenter recorded his/her name and then read the standard instructions. Assured that the subject understood them and how to rate the colors, he then presented each of the cards and recorded the assigned ratings. To ensure that the instructions were properly read and that the experimenters did not overtly communicate their expectancies to the subjects, the entire experimenter-subject interaction was monitored. A hidden microphone in the room was attached to a tape recorder in an adjacent room for this purpose. Both experimenters and subjects were unaware of this recording.

Upon completion of the color-judgment task each subject was led to Room III where the principal investigator administered the Psychological Research Survey (Adair & Fenton, 1970), designed to measure subjects' attitude toward psychological research, and the post-experimental questionnaire (Appendix H). Finally, they were given a short test of color-blindness, on which none of the subjects was found color-blind. The subjects were then debriefed about the purpose of each task and the reasons for the way they were recruited for the study.

At the conclusion of each experimenter's participation, a similar enquiry was conducted. The experimenter was asked his view of the experiment, how it was conducted, and what features about the study he noted. In

particular, the interview was designed to note if there was any suspicion as to the purpose of the study. None of the experimenters expressed suspiciousness. In conclusion, each experimenter was paid the amount of money promised and was thanked for his help.

CHAPTER III

RESULTS

An underlying assumption of this study was that the subjects who volunteered under a greater freedom of choice must have been strongly motivated to participate. This motivation could have come from, among other things, their desire to please others and/or from a positive attitude toward psychological research. Certainly it was expected that such characteristics would be stronger with these subjects than those who were less free in their choice to participate by virtue of their participation in a compulsory subject pool. In order to test the validity of this assumption, all subjects were administered the Psychological Research Survey (PRS) and the Marlowe-Crowne Social Desirability Scale (S-D) which measured their attitude toward psychological research and need for social approval, respectively.

A one-way analysis of variance for the PRS scores showed that the main effect of volunteer status was highly significant ($F = 7.01$; $df = 1, 126$; $p < .01$). As expected the volunteers had significantly more positive attitude toward psychological research than the coerced subjects. A similar analysis of the S-D scores showed that the main effect of volunteer status did not reach an acceptable level of significance ($F = 3.69$; $df = 1, 126$; $p < .06$), although the volunteers tended to show higher need for social approval than the coerced subjects. However, even without this difference in approval need, the analysis on attitudes toward psychological research indicated that the present recruitment procedures had been effective in obtaining samples of divergent attitudes.

The difference between volunteers and coerced subjects was also evident from their answers to the post-experimental questionnaire (Appendix H). For example, while they were asked to rate the perceived significance of the experiment on a 5-point scale, volunteers rated the study higher ($\overline{X}$ = 3.13) than coerced subjects ($\overline{X}$ = 2.63). Analysis of the difference in ratings indicated that volunteers had a significantly more positive view of the experiment than coerced subjects (F = 6.22; df = 1, 126; $p < .02$). Similarly, when asked why they volunteered to participate in the experiment, the typical answers of the volunteers were "To help the experimenter", "To help research", "Like participating in experiments". In contrast, the frequent answer of all coerced subjects was "To obtain credit". A few of them added "Curiosity" and "Liking for experiments", and only one subject wrote "To help research". It was also observed that the volunteers,in general,were more particular in answering each question compared to the coerced subjects. For example, when the subjects were asked what they thought about the purpose of the experiment, all volunteers attempted to give some answer to the question, whereas, 11% of the coerced subjects reported "Don't know" (χ^2 = 5.77; df = 1; $p < .02$). One of the coerced subjects, in fact, put, "How the hell should I know?" Thus, in general, there is the suggestion that the volunteers had more positive attitude toward psychological research and acted more conscientiously than the coerced subjects.

With these differences in attitudes, it was expected that there would be accompanying differences in subjects' sensitivity and response to experimental manipulations. Presumably, the more positive attitude toward research would be reflected in the volunteers' greater sensitivity and willingness to cooperate. Thus, it was hypothesized that in an ambiguous

task requiring judgments, volunteers would more often confirm the experimenter's hypothesis than coerced subjects. It was also predicted that the expectancy effect for coerced subjects would be more pronounced under heightened evaluation apprehension, because such inducement would make them highly sensitive to the experimenter's cues for adequate performance. Volunteers, being already sensitive and approval seeking, were not expected to show any significant difference in confirming experimenter expectancies under high and low evaluation apprehension. Their confirmation of experimenter expectancies were, however, predicted to be higher than that of the coerced subjects under both levels of evaluation apprehension.

To check the effectiveness of the evaluation apprehension manipulation the post-experimental ratings of subjects' concern and anxiety were compared across experimental treatments. The mean concern ratings under high and low evaluation apprehension were 3.41 and 3.48 and those for anxiety were 2.87 and 2.57, respectively. Analyses of variance of these ratings at the two levels of volunteer status and the two levels of evaluation apprehension did not show any significant differences. Thus, the post-experimental measures failed, as they also did in Johnson's study (1970), to indicate the effectiveness of this manipulation.

The dependent variable of the study was the mean ratings of colors (10 ratings per subject) which were obtained by experimenters who were led to expect either high (+5) or low (-5) average ratings from their subjects under conditions of volunteer status and evaluation apprehension. The mean rainges of the subjects obtained in each condition are shown in Table 1. To test the hypotheses, these scores were analyzed by a 2 x 2 x 2 analysis of variance for fixed effects with a harmonic mean solution for unequal n's per cell. A summary of this analysis is presented in Table 2.

TABLE 1

Mean Color-Ratings of Volunteer and Coerced Subjects with High and Low Evaluation Apprehension Tested by Experimenters with High and Low Expectancies

		Volunteer Subjects Experimenter Expectancy				Coerced Subjects Experimenter Expectancy			
		High		Low		High		Low	
Evaluation Apprehension	High	0.50		2.70		0.90	-3.50	-2.30	0.10
		2.50		-0.80		-3.70	-2.60	1.50	-3.70
		0.50		1.10		1.80	-0.40	2.60	0.70
		4.20		2.20		-2.20	-1.30	0.50	-1.00
		-2.70		-4.20		0.90	0.00	-2.10	0.60
		5.70		1.50		-1.10	3.60	0.10	-0.30
		0.50		1.50		-0.90	-3.20	-0.30	3.80
		-0.60		1.00		2.90		-0.20	0.90
		-2.30		-3.30		-0.50		2.30	
		0.20		-1.50		-3.20		-0.60	
		1.20		0.30		-0.50		0.00	
	n = 12	0.80	n = 12	0.00		2.70	n = 19	1.00	n = 20
	Mean =	0.88		0.04		-0.54		0.18	
	Low	-3.70		-2.80		3.80	0.20	3.50	-1.60
		-1.20		-3.60		3.00	2.50	2.70	-0.10
		1.20		0.30		1.10	-2.70	-1.50	-3.10
		-0.80		-2.70		-1.30	0.80	-0.30	0.50
		1.50		-2.50		3.30	-0.20	-1.10	2.50
		0.70		-0.70		-3.20	-1.20	-0.70	-2.50
		1.00		1.60		-1.10	0.90	-6.20	2.60
		-0.20		-0.80		-0.10	-5.40	0.90	0.20
		-1.60		-3.70		3.30	2.70	-0.60	
		-5.10		2.20		0.30		-0.10	
		-0.70		-1.40		1.60		0.10	
	n = 12	1.90	n = 12	1.40		4.00	n = 21	-3.60	n = 20
	Mean =	-0.58		-1.06		0.58		-0.43	

TABLE 2

Summary of the Analysis of Variance for Mean Ratings of the Colors of Volunteer and Coerced Subjects with High and Low Evaluation Apprehension Obtained by Experimenters with High and Low Expectancies

Source	df	MS	F
Volunteer Status (Vol)	1	0.50	0.10
Experimenter Expectancy (EE)	1	5.13	1.05
Evaluation Apprehension (EA)	1	8.32	1.71
Vol x EE	1	1.93	0.40
Vol x EA	1	17.73	3.64
EE x EA	1	3.80	0.78
Vol x EE x EA	1	8.23	1.69
Within Cells	120	4.87	
Total	127		

p .05 = 3.92

Volunteer Status and Experimenter Expectancy Effect

In statistical terms, the study predicted a significant interaction between volunteer status and experimenter expectancy (Hypothesis 1). That is, the difference between the ratings under high and low expectancies obtained from volunteers was expected to be significantly greater than that with coerced subjects. Although the graphical representation of the data (Fig. 1) appears to be consistent with the hypothesis, the non-significant interaction ($F < 1$; $df = 1, 120$) fails to support this prediction.

Volunteer Status, Experimenter Expectancy, and Evaluation Apprehension

The second hypothesis predicted a triple interaction among the independent variables such that experimenter expectancy effect for volunteers would be equally greater than that of coerced subjects under both high and low evaluation apprehension; whereas, coerced subjects would show greater expectancy effect under high than under low evaluation apprehension. The results of the analysis of variance show that the predicted interaction did not attain significance ($F = 1.69$; $df = 1, 120$). It is very likely that this was due in part to the fact that the evaluation apprehension had differential effects on volunteers and coerced subjects. This was indicated by an interaction between volunteer status and evaluation apprehension that approached significance ($F = 3.64$; $df = 1, 120$; $p < .06$). Since the interaction was quite large, the data were further analyzed to examine the simple effects of evaluation apprehension for volunteers and coerced subjects separately (Winer, 1962). The results of this analysis, presented in Table 3, show that contrary to the prediction, the differential effect of evaluation apprehension was significant for volunteers ($F = 5.54$; $df = 1, 120$; $p < .04$) but not for coerced subjects ($F < 1$, $df = 1, 120$). As shown in Fig. 2, the level of ratings of volunteers were consistent with the level of evaluation apprehension. With coerced subjects, an opposite effect,

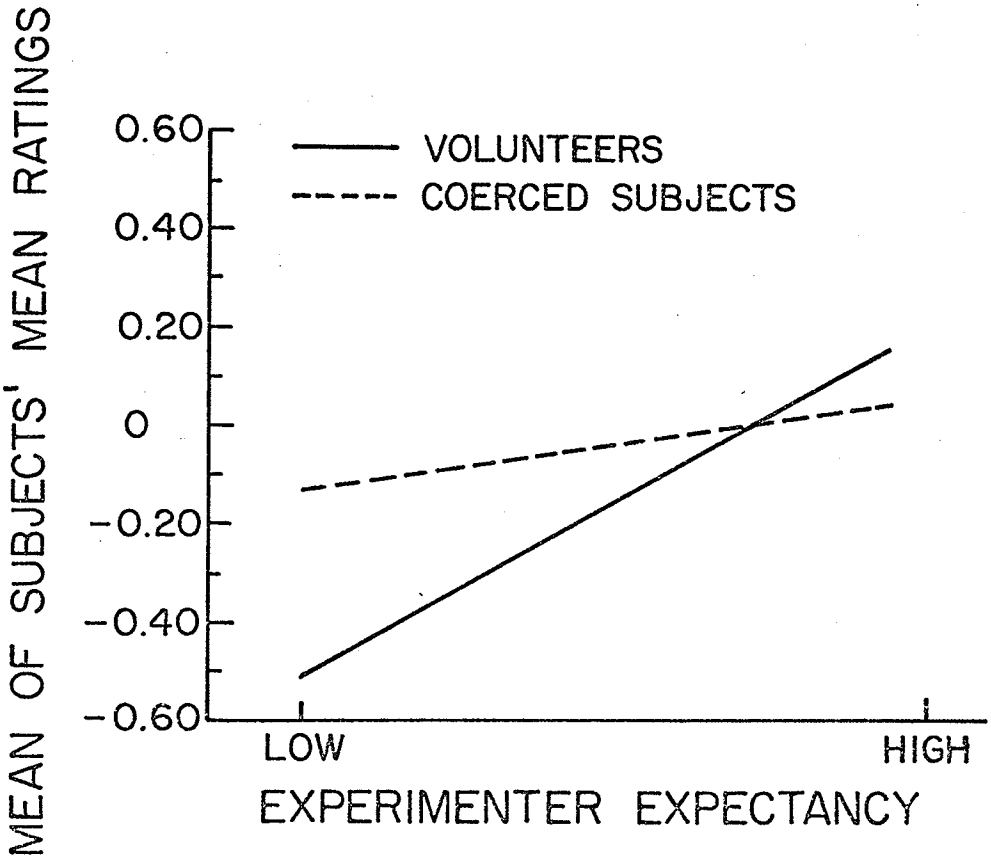


Fig. 1 Mean ratings of volunteer and coerced subjects obtained by experimenteers with high and low expectancies

TABLE 3

Summary of the Analysis of Variance for Simple Effects of High and Low Evaluation Apprehension on Volunteer and Coerced Subjects

Source	df	MS	F
Evaluation Apprehension for Volunteers	1	27.00	5.54*
Evaluation Apprehension for Coerced Subjects	1	1.25	0.25
Within Cell	120	4.87	

*$p < .04$.

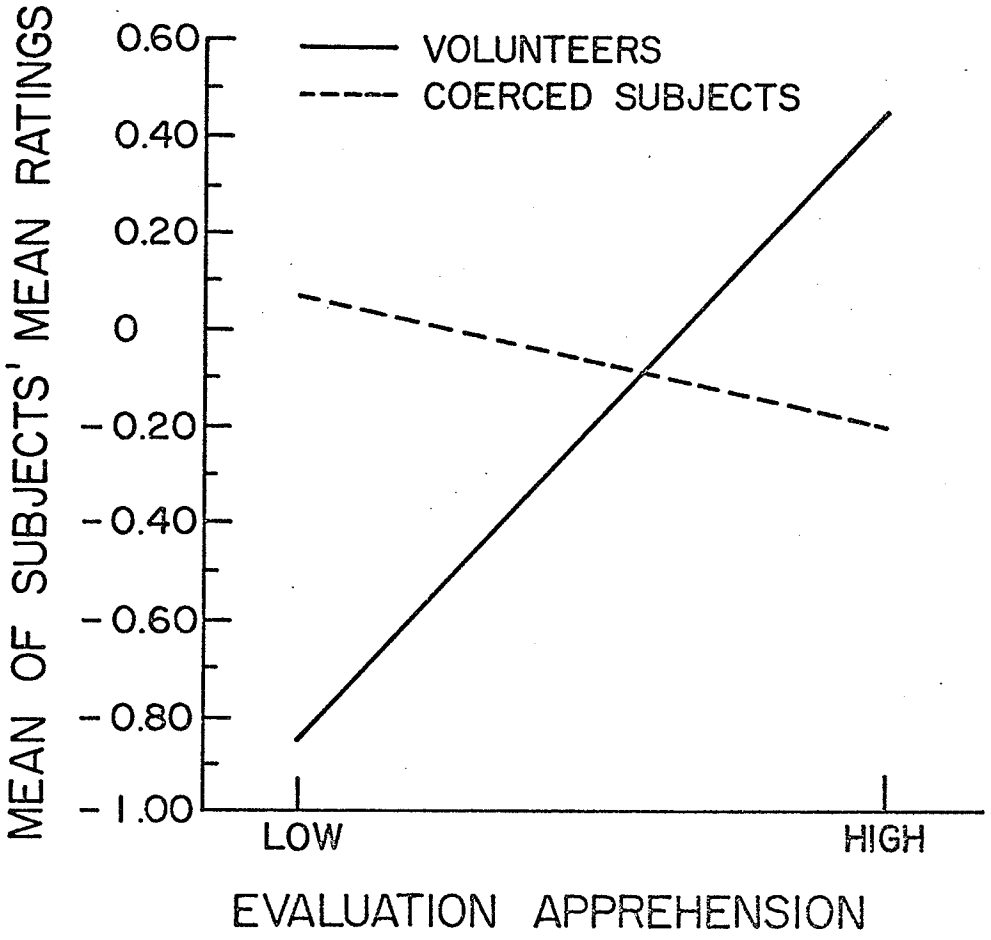


Fig. 2 Subjects' mean color-ratings under high and low evaluation apprehensions

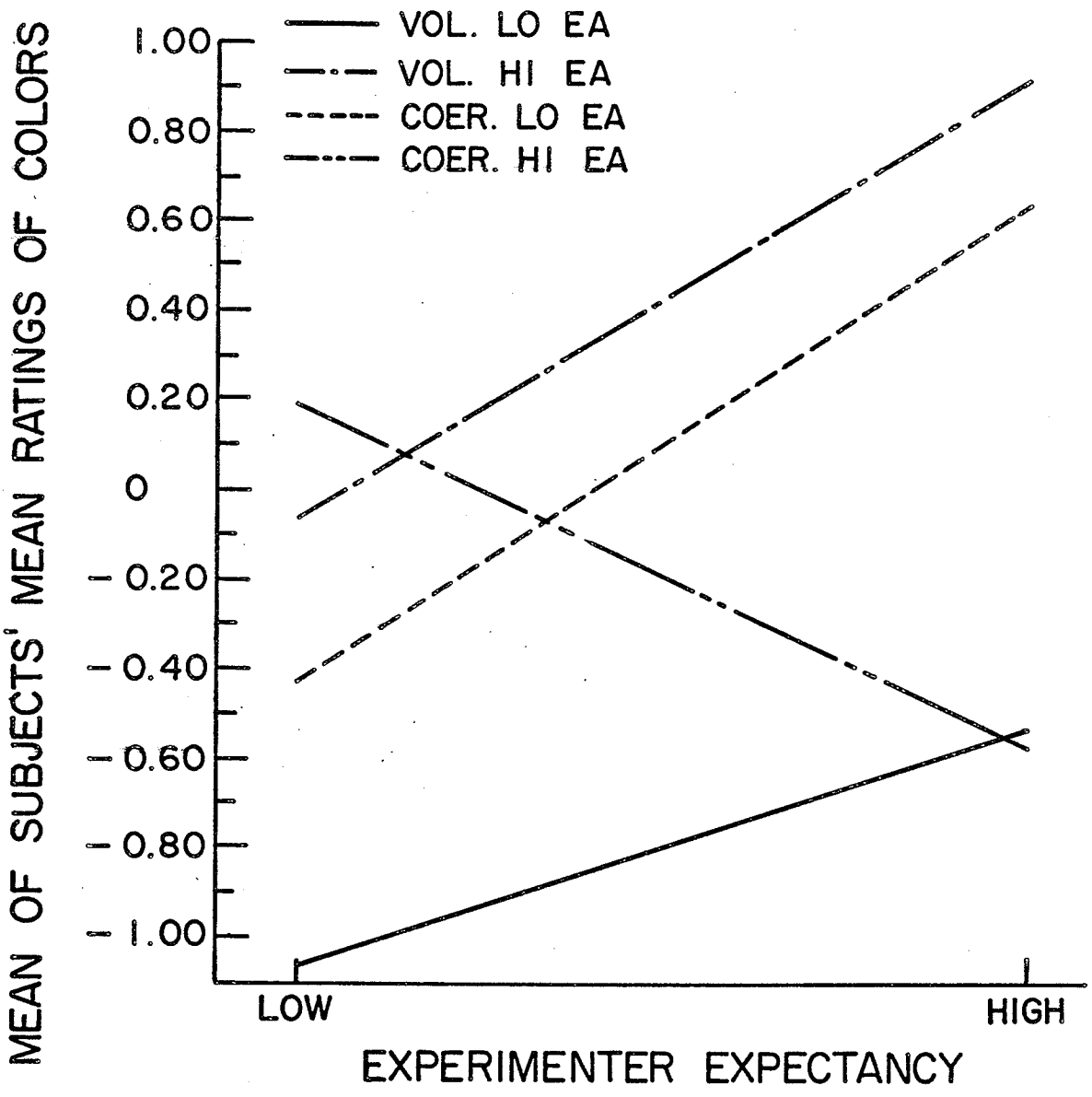


Fig. 3 Mean color-ratings of volunteer and coerced subjects under high and low evaluation apprehension obtained by experimenters with high and low expectancies

although non-significant, occurred.

Considering the non-significance of this interaction, the triple interaction among the major variables of this study, graphically represented in Fig. 3, was examined further. Examination of this figure reveals that while the ratings of volunteers were consistent with experimenters' hypothesis under both high and low evaluation apprehension, coerced subjects rated the colors in the direction of experimenters' hypotheses only under low evaluation apprehension. Contrary to the prediction, coerced subjects rated the colors opposite to experimenter expectancies while they were induced with high evaluation apprehension.

Thus, subjects' volunteer status seems to be a potential factor in determining their responses under different conditions of evaluation apprehension and experimenter expectancy. Although the interaction was not significant, it was apparent that volunteers tended to confirm experimenters' expectancies more than coerced subjects, particularly under high evaluation apprehension.

Volunteer Status, Experimenter Expectancy, Evaluation Apprehension, and Sex of Subjects

Although some of the predicted trends in the data were in the hypothesized direction, the effects may have been operative in a more complex manner with other variables. In view of the extensive prior research showing that female subjects, particularly while interacting with male experimenters, were more susceptible to the effects of experimenter expectancy (Silverman, 1968b; Johnson & Adair, 1970; Harris, 1970, 1971; Adair, 1972), it appears that sex of subjects may be a variable to consider. Since the present study used both male and female subjects and the experimenters were male, and the number of subjects falling in each cell was

not too small, the data were analyzed by a 2 x 2 x 2 x 2 factorial design for fixed effects with a harmonic mean solution for unequal n's per cell. The data are contained in Table 6 in Appendix I and the results of this analysis are presented in Table 4.

The results show that there was a significant interaction among volunteer status, experimenter expectancy, and sex of subjects ($F = 9.27$; $df = 1, 112$; $p < .004$). An examination of this interaction, graphically presented in Fig. 4, indicates that while female volunteers rated the colors consistent with experimenter's expectancies, male volunteers rated them in the opposite direction, i.e., higher under low expectancy than under high expectancy. With coerced subjects, on the other hand, the exact opposite pattern of interaction was obtained, i.e., female subjects responded opposite to experimenter expectancies while male subjects tended to confirm their expectancies. A consequence of this interaction with sex of subject presumably was to cancel out the predicted overall interaction between experimenter expectancy and volunteer status. Thus, experimenter expectancy effect was found to be a function of both volunteer status and sex of subjects.

In addition to these results, a significant interaction between evaluation apprehension and sex of subjects ($F = 4.53$; $df = 1, 112$; $p < .05$) was obtained. As shown in Fig. 5, the female subjects, in general, rated the colors consistent with the expected effects of levels of evaluation apprehension. The ratings of the male subjects were in the opposite direction. This indicates that the effects of evaluation apprehension were in the predicted direction only with female subjects but not with male subjects.

Finally, the results showed a significant overall main effect of sex of subjects ($F = 6.53$; $df = 1, 112$; $p < .02$) which is of less importance

TABLE 4

Summary of the Analysis of Variance for Mean Ratings of the Colors of Male and Female Volunteer and Coerced Subjects with High and Low Evaluation Apprehension Obtained by Experimenters with High and Low Expectancies

Source	df	MS	F
Volunteer Status (Vol)	1	0.43	0.10
Experimenter Expectancy (EE)	1	1.72	0.40
Evaluation Apprehension (EA)	1	1.16	0.27
Sex	1	28.13	6.53**
Vol x EE	1	0.00	0.00
Vol x EA	1	10.35	2.40
Vol x Sex	1	1.21	0.28
EE x EA	1	7.19	1.67
EE x Sex	1	0.00	0.00
EA x Sex	1	19.54	4.53*
Vol x EE x EA	1	8.49	1.97
Vol x EE x Sex	1	39.97	9.27***
Vol. x EA x Sex	1	3.22	0.75
EE x EA x Sex	1	5.56	1.29
Vol x EE x EA x Sex	1	0.67	0.16
Within Cells	112	4.31	
Total	127		

*$p < .05$
**$p < .02$
***$p < .004$

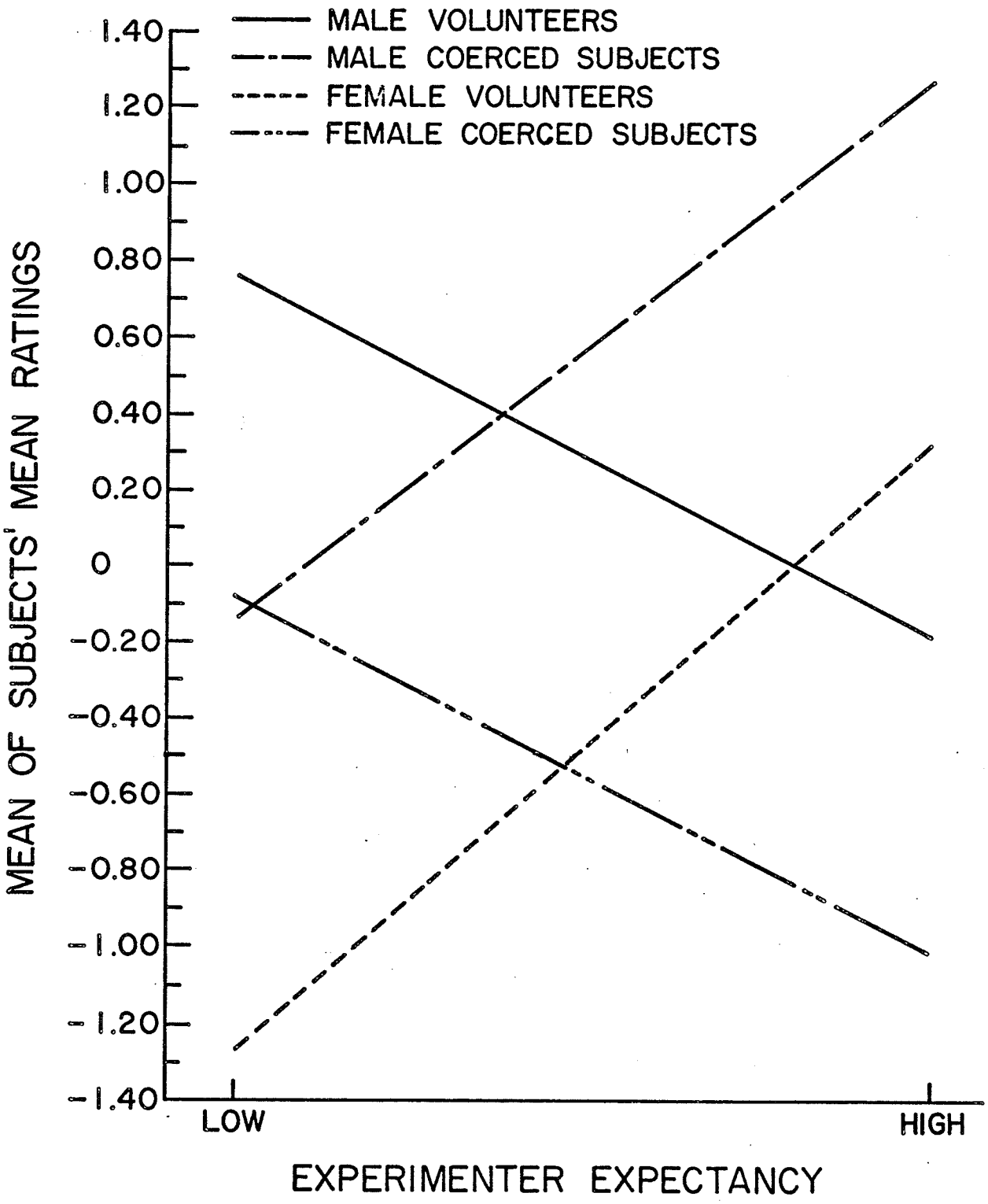


Fig. 4 Mean ratings of male and female volunteer and coerced subjects obtained by experimenters with high and low expectancies

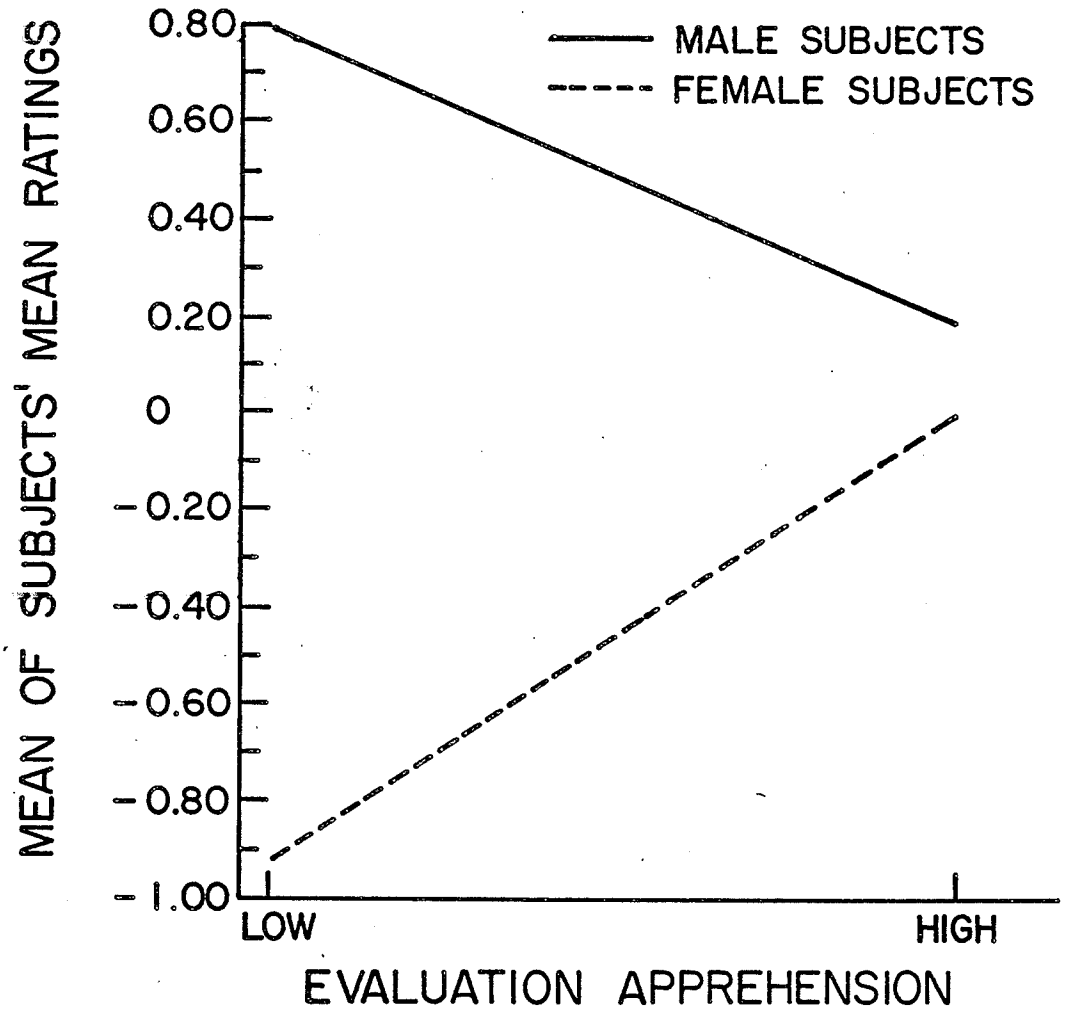


Fig. 5 Mean ratings of male and female subjects under high and low evaluation apprehension

as far as the purpose of the study is concerned. It merely adds to other information that the male subjects rated the colors more positively ($\bar{X}$ = .47) than the female subjects($\bar{X}$ = -.51).

Experimenter Expectancy Effect

The third hypothesis predicted a significant main effect of experimenter expectancy, i.e., the mean ratings of subjects under high expectancy were expected to be higher than those obtained under low expectancy. Although the mean ratings for each condition of expectancy were consistent with the prediction, the non-significant main effect of experimenter expectancy (F = 1.05; df = 1, 120) in the analysis, reported in Table 2, failed to support the hypothesis. For example, while the ratings obtained from female volunteers and male coerced subjects were consistent with experimenter expectancies, those obtained from male volunteers and female coerced subjects were exactly in the opposite direction (Fig. 4). Under the high evaluation apprehension condition, coerced subjects, in general, rated the colors opposite to experimenter expectancies (Fig. 3). Since expectancy effects varied so much as a function of subjects' sex, volunteer status, and the inducement of evaluation apprehension, no overall main effect was obtained.

Experimenter Expectancy Effect: Subjects' Attitude Toward Psychological Research and Need for Social Approval

In order to examine if experimenter expectancy effect was determined by subjects' attitude toward research and/or need for approval irrespective of their volunteer status, two analyses were made. In the first analysis, the mean color ratings of volunteer and coerced subjects at the two levels of experimenter expectancy and evaluation apprehension were split into two halves at the median according to their PRS scores, and were analyzed in a 2 x 2 x 2 x 2 factorial design for fixed effects with least

squares solution for unequal n's. The results of this analysis, presented in Table 5, failed to show any significant interaction of experimenter expectancy with subjects' attitude toward research (F = 1.53; df = 1, 112). The only effect that approached significance was the main effect of attitude toward research (F = 3.23; df = 1, 112; $p < .08$). This indicated that subjects with more positive attitudes toward research tended to rate the colors higher than those with less positive attitudes. In essence, the results failed to establish the importance of subjects' attitude toward research in the mediation of experimenter expectancy effect.

A similar analysis performed on the color-ratings obtained from subjects with high and low need for approval also did not show any significant interaction or main effects of any of these variables (Appendix I, Table 7). Thus the results failed to indicate need for approval as an important factor in determining experimenter expectancy effect.

Finally, since the effects due to sex of subjects may have been mediated by subjects' attitudes toward research, a further analysis of the PRS scores obtained by male and female subjects at the two levels of volunteer status were analyzed by a 2 x 2 analysis of variance with harmonic mean solution for unequal n's. The results of this analysis, presented in Table 8 in Appendix I, failed to indicate any significant effect of sex. Female subjects did not score higher on PRS than male subjects. Therefore, the differential responses of male and female subjects to experimenter expectancy could not be attributed to sex differences in attitude toward research. The results also failed to show any significant interaction between volunteer status and sex of subjects.

TABLE 5

Summary of the Analysis of Variance for Mean Ratings of Colors by Volunteer and Coerced subjects with High and Low Attitude Toward Psychological Research Obtained Under High and Low Experimenter Expectancies and High and Low Evaluation Apprehension.

Source	df	MS	F
Volunteer Status (Vol)	1	2.18	0.45
Experimenter Expectancy (EE)	1	6.96	1.44
Evaluation Apprehension (EA)	1	4.10	0.85
Attitude (Att)	1	15.58	3.23
Vol x EE	1	9.55	1.98
Vol x EA	1	11.72	2.43
Vol X Att	1	1.98	0.41
EE x Att	1	7.40	1.53
EE x EA	1	12.10	2.51
EA x Att	1	0.61	0.12
Vol x EE x EA	1	11.19	2.32
Vol x EE x Att	1	4.38	0.91
Vol x EA x Att	1	0.52	0.10
EE x EA x Att	1	0.19	0.04
Vol x EE x EA x Att	1	8.16	1.68
Error	112	4.83	
Total	127		

p .05 = 3.92

CHAPTER IV

DISCUSSION

This study was designed to examine the effects of subject's volunteer status and evaluation apprehension on experimenter expectancy effects. It was assumed that volunteers would be more strongly motivated to cooperate than coerced subjects. In an attempt to test the validity of this assumption, subjects' attitudes toward psychological research and need for social approval were measured. As expected, volunteers were found to have significantly more positive attitudes toward research. This was consistent with previous findings (Adair, 1970). Although not found to be significant, volunteers tended to show a higher need for approval than coereed subjects. The more positive motivation of volunteers was evident from their responses to the post-experimental questionnaire in which they assessed this experiment as significant and appeared to answer questions more conscientiously.

It was expected that with these differences in attitude and motivation toward research, volunteers would be more sensitive to the demand characteristics of the experiment and cooperative with the experimenter than coerced subjects. Specifically, it was predicted that greater experimenter expectancy effects should be found with volunteers than with coerced subjects. It was also predicted that volunteers would show greater experimenter expectancy effects under both high and low evaluation apprehension, whereas, coerced subjects would confirm the experimenter's hypothesis more under high evaluation apprehension.

Finally, a significant experimenter expectancy effect was expected across all treatment conditions.

The results failed to support these hypotheses. Although the data obtained by the experimenters were in the direction of their expectations, the study failed to show a significant expectancy effect either overall or in complex interactions with volunteer status and evaluation apprehension. A possible reason for this could be the experimental task. Although the color-judgment task was selected because it had been previously found to be effective in demonstrating the phenomenon (Shames, 1971), the absence of bias effect in the present study together with its limited prior use suggest that the task may not be as susceptible to experimenter expectancy effect as originally thought. Indeed, the similarity of this task to the person perception task in both rating scales and general procedure suggests that it may suffer from some of the same inconsistencies of the latter task in demonstrating the experimenter expectancy effect (Rosenthal,1969).

Another possibility is that the experimenters' expectancies were not effectively communicated to the subjects. On the one hand, the experimenters might not have been sufficiently biased. Although it was the impression of the principal investigator in post-experimental interviews that experimenters believed what they were told, there was no way to assess whether or the degree to which they really accepted the hypothesis. For example, experimenters may have merely been ingratiating to the principal investigator during these interviews. If they were not strongly convinced of the hypothesis, they would have failed to emit the appropriate cues to the subjects with consistency. Under these conditions, even volunteers, who were willing to cooperate, could not possibly produce the

data systematically in the direction of the induced expectancies.

On the other hand, the subjects, for some reason, might not have been sensitive to whatever cues were transmitted by the experimenters. Under the present experimental set-up subjects' attention may have been directed more toward the manipulation of other variables such as evaluation apprehension than toward the expectancy cues. The description about the purpose of the study which they read before performing the color-judgment task might have been a sufficiently strong cue to subjects to communicate how they should respond that they might not have needed to look to the experimenter for additional cues. For example, while they read the description intended to induce high evaluation apprehension, some subjects might have decided to cooperate with the experimenter whereas others did not intend to do so. The data provide some support for this interpretation. Volunteers rated the colors significantly higher under the high than under the low evaluation apprehension condition. Coerced subjects, on the other hand, responded significantly in the opposite manner to volunteers. This differential reaction of volunteers and coerced subjects may suggest that perhaps in the present experimental set-up the manipulation of evaluation apprehension worked more as a demand characteristic suggesting the experimental hypothesis and the way the subjects should respond rather than a source of motivation to look for the experimenter's cues for 'correct' responding. From this view, then, all subjects, particularly volunteers who had a greater tendency to cooperate, would have confirmed the experimenter's expectancy more often if the treatment of evaluation apprehension had not encouraged them to

develop a set and thus become less attentive to the experimenter's cues.

Although the predicted interactions were not significant, the data indicated some differences between volunteers and coerced subjects in response to the experimenter's expectancy under different levels of evaluation apprehension. For example, unlike volunteers, coerced subjects tended to rate colors opposite to the experimenter's expectancies under high evaluation apprehension. This may be interpreted as an intentional negativistic response or as Masling (1966) has called it, a "screw you effect". A possible explanation for this unpredicted effect may be that these subjects, 'coerced' by a course requirement to participate in experiments, may include many who would otherwise be unwilling to serve as subjects and who therefore may have had a covert resentment against compulsory participation. There is some indication of this in their attitude toward research which was significantly less positive than that of volunteers. In addition, the evaluation apprehension treatment might have made them aware of the influence attempt by the experimenter whom they probably had associated with their coerced status. With these poor attitudes, coerced subjects might have attempted to ruin the experiment by rating the colors opposite to the perceived expectancies of the experimenter. In contrast, volunteers, having a more positive attitude toward participating in research, tended to confirm the perceived hypothesis. Although the experimenter expectancy effect did not differ significantly with volunteer status and evaluation apprehension, these data suggest that evaluation apprehension may have differential effects on the gross level of responses of volunteers and coerced subjects.

In an attempt to investigate the variables that might have contributed to the non-significance of the predicted interactions among experimenter expectancy, volunteer status and evaluation apprehension, the data were further analyzed according to sex of subjects. From this post-hoc analysis several observations were made. First, the experimenter expectancy effect was found to vary significantly as a function of subjects' sex and volunteer status. This triple interaction indicated that within the volunteer group female subjects confirmed the experimenter's hypothesis more than males. This finding was consistent with the results of previous studies (Silverman, 1968b; Johnson & Adair, 1970; Adair, 1972) where female subjects tested by male experimenters confirmed the experimenter's hypothesis more than male subjects tested by male experimenters. As suggested in these studies, male experimenters were perhaps more efficient in transmitting the appropriate expectancy cues to female subjects as they were likely to be more interested in the members of the opposite sex. It could be that the experimenters, all of whom were males, paid more attention to female subjects and wanted them to perform well (Rosenthal, 1966). As a result, they might have been more efficient in giving these subjects the right cues. To male subjects, on the other hand, the experimenters might have transmitted the inappropriate cues as they probably were less concerned and less attentive to the members of the same sex. This differential treatment together with the positive attitudes of volunteers toward research may have resulted in a greater confirmation of their expectations by female subjects and a "boomerang effect" (Silverman, 1968a) with males.

In the coerced group, on the other hand, the interaction between experimenter expectancy and sex of subjects was reversed. That is, female subjects rated the colors opposite to the experimenter's hypothesis, whereas, male subjects confirmed them. This reversal may be explained in terms of the differential experimenter-subject interaction due to sex along with the poor attitude of coerced subjects toward participating in experiments discussed earlier. Since this experiment was conducted at the end of the academic year, it is very likely that the coerced samples included more non-volunteers than usual (Adair, 1970). It could be that although these coerced subjects, with poor attitudes toward research, received the same cues from the experimenter (which supposedly were more accurately transmitted to females than males) as did the volunteers, they might have wanted to ruin the experiment by responding opposite to the perceived hypothesis.

It should be noted in this context that the time of the year of this study may have resulted in both samples of volunteers and non-volunteers being quite extreme in comparison with usual samples of subjects. It is very likely that the coerced subjects, who had not completed their requirements for the experimental credit until this time of the year, included more non-volunteers and hence had poorer attitudes toward research than usual. Similarly, subjects in the volunteer group may have been an extreme type of volunteer since they offered their help in spite of being busy with final examinations and having already completed their requirements. This extremeness of volunteer and coerced samples suggests the need for caution in generalizing the results.

Although expectancy effects did not significantly relate to evaluation apprehension, the responses of male and female subjects were

found to be differentially affected by the conditions of evaluation apprehension. In general, females rated the colors significantly higher than males under heightened evaluation apprehension. This finding may have implications for the interaction between volunteer status and evaluation apprehension obtained in the original analysis. It should be noted that unlike the coerced group, almost two-thirds of volunteers were females. Therefore, it may be reasoned that the higher color-ratings of volunteers under high evaluation apprehension was due to sex of subjects which was a correlate of volunteering.

Implications and Need for Further Research

The study has indicated the complexity of relationship between volunteer status and sex of subjects in determining the experimenter expectancy effects. Volunteer status and sex of subjects may be considered significant determinants of expectancy effects, however, only in appropriate combinations of subject and experimenter sex. Although it was speculated why female volunteers and male coerced subjects confirmed the experimenter's hypothesis while male volunteers and female coerced subjects did not, it is subject to further research to reveal the exact processes determining these effects. For example, a thorough examination of the video tape-recordings of verbal and non-verbal behaviors of the experimenters and of volunteer and non-volunteer subjects of the same and opposite sex may suggest the variables underlying these complex interactions.

It is important to note that female subjects confimred the experimenter's hypothesis only when they volunteered. This finding has important implications for research in which subjects are recruited on a

voluntary basis without restrictions on sex sampling. Since females are more likely to volunteer (Rosenthal & Rosnow, 1969), the results of these studies may be biased in the direction of the experimenter's hypothesis. In addition, the observation that female subjects, tested by male experimenters, are more influenced by experimenter expectancies may be limited to the females who volunteered. However, further research is necessary to make generalizations about the sex differences in this regard.

This study also indicated the importance of volunteer status in determining the effects of evaluation apprehension. The finding that volunteers had a tendency to confirm experimenter expectancies while coerced subjects showed a "boomerang effect" under high evaluation apprehension suggests that the experiments in which evaluation apprehension is a potential variable the use of volunteer subjects may bias the results toward the experimenter's hypothesis. Apart from volunteer status, sex was found to be an important factor determining the effects of evaluation apprehension. The higher ratings of females than males under high evaluation apprehension suggests that female subjects may be more sensitive to the conditions of evaluation apprehension. Both of these observations have implications for the evaluation apprehension research of Rosenberg and his associates (Rosenberg, 1969). The effects of evaluation apprehension appear to be enhanced by using either volunteers or female subjects and any research manipulating evaluation apprehension or in which evaluation apprehension may be internally aroused may show inconsistent results depending upon the sample selected. For example, this finding may account for the inconsistencies in experimenter expectancy studies in which the effects of evaluation apprehension have been manipulated (Minor, 1970; Johnson, 1970b).

Considering that the manipulation of evaluation apprehension was a prominent feature of the present experimental set-up which might have prevented subjects' attention to the expectancy cues by inducing a set as to how they should respond, it is suggested that in subsequent replications a control group (i.e., a group without receiving prior description about the study) he added. Such a group would provide a baseline to compare subjects' responses to expectancy cues without the presence of other demand characteristics with their responses in other conditions in which the manipulation of evaluation apprehension was prominent. This would enable a clearer examination of the contribution to the experimenter expectancy phenomenon of both volunteer status and evaluation apprehension.

Since this experiment was conducted at the end of the academic year, more non-volunteers may have been included in the coerced group, and this would have enhanced the difference between volunteers and coerced subjects. Thus, it may be important to replicate the study earlier in the year to assess the generality of the findings with volunteers and coerced subjects. In addition, variation of recruitment procedures for obtaining both volunteers and coerced subjects is required to further understand subjects' differential reactions to the demand characteristics of the experimental situation. Similarly, the study should be replicated with other experimental tasks to indicate the differences in the results due to the task. It is hoped that these suggested studies will provide information about the effects of subjects' volunteer status as well as the conditions under which experimenter expectancy effect may be obtained.

CHAPTER V

SUMMARY AND CONCLUSION

This study was designed to examine the effects of volunteer status and evaluation apprehension on experimenter expectancy effect. Two groups of subjects, volunteers and coerced subjects, were recruited from different sections of the Introductory Psychology course which required every student to obtain five hours of experimental credit. Volunteers were 16 male and 32 female students who served as subjects on a completely voluntary basis, i.e., with the understanding that there would be no experimental credit for participation. Coerced subjects, on the other hand, were 39 male and 41 female students who participated to obtain one hour of credit.

Prior to testing on the experimental task, one half of the subjects from each group was induced with high and the other half with low evaluation apprehension by two sets of printed descriptions about the purpose of the study. The experimental task consisted of ten chromatic colors which were to be rated in terms of color-richness on a 20-point scale. Subjects were tested on this task by one of the 18 male experimenters who were previously biased to expect either a high (+5) or low (-5) average ratings from their subjects. Furthermore, their attitude toward psychological research and need for approval were assessed by administering the Psychology Research Survey (Adair & Fenton, 1970) and Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1964) before and after the experimental task, respectively.

As expected, volunteers scored higher on both measures, indicating a significantly more positive attitude toward research and a tendency to a higher need for approval than coerced subjects. Assuming that these differences in attitudes and needs would affect their reactions to the cues of the experiment, it was hypothesized that experimenter expectancy effect, i.e., the difference between subjects' mean ratings under high and low expectancies, would be greater for volunteers than for coerced subjects. Secondly, it was expected that the volunteers' expectancy effect would be equally greater under high and low evaluation apprehension. Coerced subjects, on the other hand, were predicted to show greater expectancy effect under high evaluation apprehension than under the low evaluation apprehension condition. Finally, a significant experimenter expectancy effect was predicted across all treatment conditions.

The results of an analysis of variance of subjects' mean ratings of colors failed to support these hypotheses. Although subjects' responses on the experimental task were in the predicted direction, volunteers did not show significantly greater expectancy effect than coerced subjects. Similarly the expectancy effect of volunteers and coerced subjects did not differ significantly as a function of evaluation apprehension. It was found, however, that the responses of volunteers were more enhanced than those of coerced subjects under high evaluation apprehension. This was indicated by a significant simple effect of evaluation apprehension with volunteers. Finally, the analysis failed to show a significant overall experimenter expectancy effect.

Post-hoc analysis according to sex of subjects revealed a signifi-

cant triple interaction among experimenter expectancy, sex, and volunteer status. Female volunteers and male coerced subjects were found to have confirmed the experimenter's hypothesis more than male volunteers and female coerced subjects. This indicated that sex of subjects may combine with volunteer status in a complex way to determine experimenter expectancy effect. There was also a significant interaction between sex of subjects and evaluation apprehension which indicated that female subjects, in general, were more positively affected by evaluation apprehension than male subjects. Further post-hoc analysis of these data according to subjects' level of attitude toward psychological research and need for approval failed to show any additional significant effects.

On the basis of these findings, it was concluded that a number of subjects' characteristics may combine in a complex way to determine their responses in an experimental situation. Whether the experimenter's expectancies would be confirmed or not may be a function of subjects' volunteer status and sex. From the implications of this study it was thought that in experiments involving an unstructured task, e.g., color-judgment, the expectancies of a male experimenter may be communicated more effectively to the subjects of the opposite sex. If these subjects happen to be volunteers, confirmation of his hypothesis may be obtained. The effects of evaluation apprehension may also be enhanced if subjects

are volunteers and females. Further research was suggested to make a thorough examination of these variables.

REFERENCES

Adair, J.G. Pre-experimental attitude toward psychology as a determinant of subject behavior. Paper presented in the symposium entitled "Methodological problems in research with human subjects" at the annual meeting of the Canadian Psychological Association, Winnipeg, Canada, May, 1970.

Adair, J.G. Demand characteristics or conformity: Suspiciousness of deception and experimenter bias in conformity research. Canadian Journal of Behavioural Science, 1972, 4, 238-248.

Adair, J.G. The human subject: The social psychology of the psychological experiment. Boston: Little, Brown, 1973. In press.

Adair, J.G., & Fenton, D.P. Subject's attitude toward psychology as a determinant of experimental results. Canadian Journal of Behavioural Science, 1970, 3, 268-275.

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

Beez, W.V. Influence of biased psychological reports on teacher behavior and pupil performance. Proceedings of the 76th Annual Convention of the American Psychological Association, 1968, 3, 605-608.

Black, R.W., Schumpert, J., & Welch, F.A. A partial reinforcement extinction effect in perceptual-motor performance: Coerced versus volunteer subject populations. Journal of Experimental Psychology, 1972, 92, 143-145.

Blake, B.F., & Heslin, R. Evaluation apprehension and subject bias in experiments. Journal of Experimental Research in Personality, 1971, 5, 57-63.

Cox, D.E., & Sipprelle, C.N. Coercion in participation as a research subject. American Psychologist, 1971, 26, 726-728.

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

Duncan, S., Rosenberg, M., & Finkelstein, J. The paralanguage of experimenter bias. Sociometry, 1969, 32, 207-219.

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

Friedman, N. The social nature of psychological research. New York: Basic Books, 1967.

Goldstein, J.H., Rosnow, R.L., Goldstadt, B., & Suls, J.M. The good subject in verbal operant conditioning. Unpublished manuscript, Temple University, 1970.

Harris, S. Influence of subject and experimenter sex in psychological research. Journal of Consulting and Clinical Psychology, 1971, 37, 291-294.

Harris, S., & Masling, J. Examiner sex, and Rorschach productivity. Journal of Consulting and Clinical Psychology, 1970, 34, 60-63.

Horowitz, I.A. Effects of volunteering, fear arousal, and number of communications on attitude change. Journal of Personality and Social Psychology, 1969, 11, 34-37.

Johnson, R.W. Subject response as affected by experimenter expectancy, sex of experimenter, and verbal reinforcement. Canadian Journal of Behavioural Science, 1970, 2, 60-66. (a).

Johnson, R.W. Inducement of expectancy and set of subjects as determinants of subjects' responses in experimenter expectancy research. Unpublished doctoral dissertation, University of Manitoba, 1970.(b)

Johnson, R.W., & Adair, J.G. The effects of systematic recording error vs. experimenter bias on latency of word association. Journal of Experimental Research in Personality, 1970, 4, 270-275.

Jung, J. Current practices and problems in the use of college students for psychological research. Canadian Psychologist, 1969, 10, 280-290.

Kennedy, J.J., Cook, P.A., & Brewer, R.R. An examination of the effects of three selected experimenter variables in verbal conditioning research. Journal of Social Psychology, 1970, 81, 167-175.

Laszlo, J.P., & Rosenthal, R. Subject dogmatism, experimenter status and experimenter expectancy effects. Unpublished manuscript. Harvard University, 1967.

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

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

Masling, J. Role related behavior of the subject and psychologist and its effects upon psychological data. In D.L. Levine (Ed.), Nebraska Symposium on Motivation, Lincoln, Nebraska: University of Nebraska Press, 1966. Pp. 67-103.

McDavid, J.W. Approval-seeking motivation and the volunteer subject. Journal of Personality and Social Psychology, 1965, 2, 115-117.

Minor, M.W. Experimenter-expectancy effect as a function of evaluation apprehension. Journal of Personality and Social Psychology, 1970, 15, 326-332.

Myers, J.L. Fundamentals of experimental design. Boston: Allyn & Bacon, 1966.

Orne, M. On social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. American Psychologist, 1962, 11, 776-783.

Orne, M. Demand characteristics and the concept of quasi-controls. In R. Rosenthal & R.L. Rosnow (Eds.), Artifact in behavioral research. New York: Academic Press, 1969.

Pflugrath, J. Examiner influence in a group testing situation with particular reference to examiner bias. Unpublished Master's thesis. University of North Dakota, 1962.

Raffetto, A.M. Experimenter effects on subjects' reported hallucinatory experiences under visual and auditory deprivation. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, May, 1968.

Riecken, H.W. A program for research on experiments in social psychology. In N.F. Washburn (Ed.), Decisions, values, and groups. New York: Pergamon Press, 1962. Vol. II, 25-41.

Rosenberg, M.J. When dissonance fails: On eliminating evaluation apprehension from attitude measurement. Journal of Personality and Social Psychology, 1965, 1, 28-42.

Rosenberg, M.J. The conditions and consequences of evaluation apprehen-

sion. In R. Rosenthal & R. L. Rosnow (Eds.), Artifact in Behavioral research. New York: Academic Press, 1969. Pp. 279-349.

Rosenthal, R. Experimenter effects in Behavioral research. New York: Appleton-Century-Crofts, 1966.

Rosenthal, R. Interpersonal expectations: Effects of experimenter's hypothesis. In R. Rosenthal and R. L. Rosnow (Eds.), Artifact in behavioral research. New York: Academic Press, 1969, Pp. 182-277.

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

Rosenthal, R., & Fode, K.L. Three experiments in experimenter bias. Psychological Reports, 1963, 12, 491-511. (b)

Rosenthal, R., & Jacobson, L. Pygmalion in the classroom: Teacher expectation and pupils' intellectual development. New York: Holt, Rinehart, & Winston, 1968.

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

Rosenthal, R., & Rosnow, R.L. The volunteer subject. In R. Rosenthal & R. L. Rosnow (Eds.), Artifact in behavioral research. New York: Academic Press, 1969. Pp. 182-277.

Rosnow, R.L., & Rosenthal, R. Volunteer subjects and results of opinion change studies. Psychological Reports, 1966, 19, 1183-1187.

Rosnow, R.L., & Rosenthal, R. Volunteer effects in behavioral research. New directions in psychology, Vol. 4. New York: Holt Rinehart

and Winston, 1970, 213-277.

Rosnow, R.L., Rosenthal, R., McConchie, R.M., & Arms, R.L. Volunteer effects in experimental outcomes. Educational and Psychological Measurement, 1969, 29, 825-846.

Rosnow, R.L., & Suls, J.M. Reactive effects of pre-testing in attitude research. Journal of Personality and Social Psychology, 1970, 15, 338-343.

Shames, M.L. The ambiguity model: A dimensional analysis of experimenter expectancy effects. Unpublished doctoral dissertation, University of Manitoba, 1971.

Silverman, I. Role-related behavior of subjects in laboratory studies of attitude change. Journal of Personality and Social Psychology, 1968, 8, 343-348. (a)

Silverman, I. The effects of experimenter outcome expectancy on latency of word association. Journal of Clinical Psychology, 1968, 24, 60-63. (b)

Wartenberg-Ekran, U. The effects of experimenter knowledge of a subject's scholastic standing on the performance of a reasoning task. Unpublished Master's thesis, Marquette University, 1962.

Winer, B.J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

APPENDICES

Appendix A

Request for Volunteering

"I have been doing an experiment for my thesis for which I need some volunteers. Actually, I have already tested the number of subjects that had been assigned to me. Now I feel I should test some more subjects. But you see, I cannot get them by displaying a sign-up booklet because it is not possible for me to give any credit for participation. That is why I am requesting your help. Would you like to volunteer? If you do, please sign one of these blank sheets. I would appreciate your help."

Appendix B

Color-Judgment Scale

Extremely Color Poor			Very Color Poor			Moderately Color Poor			Moderately Color Rich			Very Color Rich			Extremely Color Rich		
-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10

Appendix C

Instructions to the Subjects

"I am going to read you some instructions. I am not permitted to say anything which is not in the instructions. O.K.?

We are in the process of developing a battery of human engineering tests, one of which entails the judgment of color. The test is designed largely to assess your perception of color. I will show you a series of cards and for each one I want you to assess the 'richness' or 'poorness' of color.

To help you make exact judgments you are to use this rating scale in front of you. As you can see, that scale runs from -10 to +10. A rating of -10 means that you judge the particular color to be extremely color poor while a rating of +10 means that you judge the color to be extremely color rich. A rating of -1 means that you judge the color to be moderately color poor while a rating of +1 means that you assess the color as being moderately color rich.

You are to rate each color as accurately as you perceive it. Just tell me the rating you assign to each color on the cards I will show you. All ready? Here is the first card."

Appendix D

Recording Sheet for Color-Ratings

Subject

Card	Rating
#1	
#2	
#3	
#4	
#5	
#6	
#7	
#8	
#9	
#10	

Experimenter .....Mean Rating .....

Appendix E

Inducement of Experimenter Expectancy

"Now, let me give you a brief description of the study. The experiment is concerned with developing a battery of human engineering tests. Perhaps you know that color can be reliably related to people's affect or emotions. For example, in Rorschach Ink Blot Tests which, by the way, are often in color (demonstrated), it has been found that people of different emotional characteristics perceive colors differently as to their richness, brightness, etc. In an attempt to measure this relationship in a more scientific manner we devised this test of rating colors on this 20-point scale (demonstrated). Previously, in our pilot study, we attempted to test the validity of this test with people to whom we had administered the Guilford-Zimmerman Temperament Survey and found that those who are highly emotional rated the colors more positively, i.e., +5 on the average, than those who are less emotional whose ratings were negative, i.e., -5 on the average.

The purpose of this study is to replicate this finding on a large sample of subjects. The subjects have been divided into high and low emotionality groups based on their scores on Guilford-Zimmerman Temperament Survey administered in the beginning of this session. Our hypothesis is that the high emotional subjects will rate the colors more positively (say +5) than the low emotional subjects who will rate them negatively (say -5). You will test subjects from both groups. This is the list of subjects you are going to test. By the way, the names marked "H" in

pencil are those who belong to the high emotionality group and those marked "L" belong to the low emotionality group."

Appendix F

High Evaluation Apprehension Communication

Today you will be helping us to collect some data for psychological research. Shortly you will be assigned to an experimenter who will explain the task to you. In order to make participating more informative and meaningful for you, we will give a brief description of the purpose of the study.

We are interested in human engineering, particularly color perception. More specifically, we want to find the factors which increase or decrease the accuracy of an individual's judgment of color. Prior research by ourselves and others indicate that poor judgment of color is associated with psychopathology. That is, apart from any physical defect, viz., color blindness, people who are not able to accurately judge colors are psychologically maladjusted. Much of our initial research in this area indicates that on the basis of performance on this task, we can pick out from a college population those students who could be judged clinically maladjusted. Morgan and Provino (1963), for example, report that in a college setting, the color-judgment test could make rather subtle discriminations between varying degrees of emotional maladjustment and normalcy. The purpose of today's experiment, therefore, is to replicate the previous results, and thus to test further the generality of the finding that people who cannot accurately judge colors tend to be psychologically maladjusted.

Appendix G

Low Evaluation Apprehension Communication

Today you will be helping us to collect some preliminary data which we will use in setting up a subsequent research project. Shortly, you will be assigned to an experimenter who will explain the task to you. In order to make participating more informative and meaningful for you, we will give you a brief description of the purpose of the study.

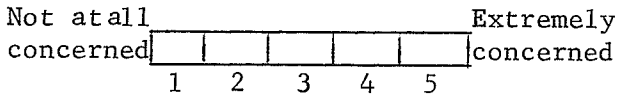
We are interested in developing a test of human engineering, particularly color perception. More specifically, we want to find the factors (e.g., fatigue, practice, etc.) which increase or decrease the accuracy of an individual's judgment of color. Before we can investigate these different factors, however, we have to know how people perceive the feelings and experiences of others when these experimental factors are not present. That is, we need a control, or standardization group to use as a baseline against which we can judge the effects that experimental factors have on color perception. This is the reason for your participation today. We intend to average the performance of all these students participating today, so that we will have a measure of how subjects perform on the task when such experimental variables as fatigue and prior practice are not present. This information will allow us to judge the effects which our experimental variables have when they are used with a subsequent group of students. In other words, today's group will help us to find out how subjects typically perform on the task. Later we can use the data we receive here to judge the performances of subsequent experimental groups of subjects.

Appendix H

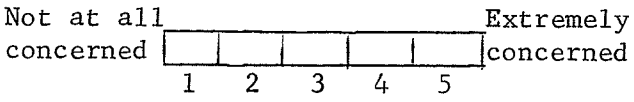
Post-Experimental Questionnaire

On the following pages are questions regarding the experiment in which you have just participated. The questions are intended to obtain your thoughts, feelings, and understandings about different aspects of the experiment so that in the future we may plan our experiments in a more scientific manner. So, please give your frank and honest opinion. Answer each question in numerical order. DO NOT look back to your previous answers.

1. Every experiment has to have some purpose. What do you think the purpose of this experiment was?
2. What aspects of the experiment led you to think so?
3. What made you volunteer for the experiment? Please specify.
4. How concerned were you with your performance on the color-judgment test? Check one.



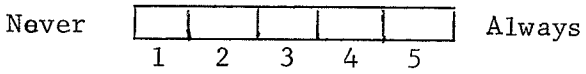
5. How did the written instructions explaining the purpose of the experiment affect you as you were about to take the color-judgment test? Check one.



6. What do you think the purpose of that written instruction was?

7. Did you look at the experimenter while rating the colors?

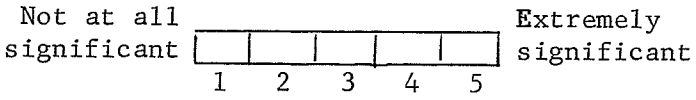
   If yes, how often?



8. Did the experimenter do anything to help you perform better on the color-judgment test? If yes, what?

9. What purpose did you think was served by the questionnaires before and after the color-judgment task?

10. Finally, how significant did the experiment appear to you?



Appendix I

TABLE 6

Mean Color-Ratings of Male and Female Volunteer and Coerced Subjects with High and Low Evaluation Apprehension Tested by Experimenters with High and Low Expectancies.

	Volunteer Subjects				Coerced Subjects			
	Experimenter Expectancy				Experimenter Expectancy			
	High		Low		High		Low	
	Male	Female	Male	Female	Male	Female	Male	Female
High	0.50	2.50	2.70	-0.80	-3.70	0.90	-2.30	2.60
	-2.30	0.50	2.20	1.10	1.80	-2.20	1.50	0.50
	0.20	4.20	1.50	-4.20	0.90	-0.90	0.10	-2.10
		-2.70	-3.30	1.50	-1.10	-0.50	-0.30	-0.20
		5.70	0.30	1.00	2.90	-3.20	2.30	1.00
		0.50		-1.50	-0.50	-3.50	-0.60	-3.70
		-0.60		0.00	2.70	-0.40	0.00	0.70
		1.20			-2.60	-1.30	0.10	-1.00
		0.80			3.60	0.00		-0.30
	n = 3	n = 9	n = 5	n = 7	-3.20		0.60	3.80
							0.90	
					n = 10	n = 9	n = 10	n = 10
Low	-1.20	-3.70	0.30	-2.80	3.80	1.10	3.50	-0.30
	1.00	1.20	1.60	-3.60	3.00	-1.30	2.70	-1.10
	-1.60	-0.80	-0.80	-2.70	3.30	-3.20	-1.50	-6.20
	1.90	1.50	2.20	-2.50	3.30	-1.10	-0.70	0.90
		0.70		-0.70	1.60	-0.10	-0.60	-0.10
		-0.20		-3.70	4.00	0.30	-3.60	-0.10
		-5.10		-1.40	0.20	-2.70	-3.10	-1.60
		-0.70		1.40	2.50	0.80	0.50	-0.10
					0.90	-0.20	-2.50	2.50
						-1.20	0.20	2.60
	n = 4	n = 8	n = 4	n = 8		-5.40		
					n = 9	2.70		
						n = 12	n = 10	n = 10

TABLE 7

Summary of the Analysis of Variance for Mean Ratings of Colors by Volunteer and Coerced Subjects with High and Low Need For Approval Obtained Under High and Low Experimenter Expectancies and High and Low Evaluation Apprehension.

Source	df	MS	F
Volunteer Status (Vol)	1	1.16	0.24
Experimenter Expectancy (EE)	1	3.47	0.71
Evaluation Apprehension (EA)	1	3.18	0.65
Need Approval (NA)	1	6.67	1.36
Vol x EE	1	3.19	0.65
Vol x EA	1	22.29	4.54*
Vol x NA	1	1.46	0.29
EE x NA	1	7.62	1.55
EE x EA	1	12.83	2.61
EA x NA	1	0.39	0.07
Vol x EE x EA	1	5.03	1.02
Vol x EE x NA	1	0.03	0.00
Vol x EA x NA	1	3.99	0.81
EE x EA x NA	1	3.19	0.65
Vol x EE x EA x NA	1	10.27	2.09
Error	112	4.91	
Total	127		

* $p < .05$

TABLE 8

Summary of the Analysis of Variance for the PRS Scores of Male and Female Volunteer and Coerced Subjects

Source	df	MS	F
Volunteer status	1	3224.55	6.89*
sex	1	157.06	0.34
Vol x sex	1	51.66	0.11
Within Cells	124	468.09	
Total	127		

* $p < .01$