4-1-1971

Relationship of Subjects' Preexperiment Attitudes Toward Psychology and Experimenter Prestige

Sam Quick
Western Kentucky University

Follow this and additional works at: http://digitalcommons.wku.edu/theses
Part of the Psychology Commons

Recommended Citation
http://digitalcommons.wku.edu/theses/1007

This Thesis is brought to you for free and open access by TopSCHOLAR*. It has been accepted for inclusion in Masters Theses & Specialist Projects by an authorized administrator of TopSCHOLAR*. For more information, please contact topscholar@wku.edu.
RELATIONSHIP OF SUBJECTS' PREEXPERIMENT
ATTITUDES TOWARD PSYCHOLOGY AND
EXPERIMENTER PRESTIGE

A Thesis
Presented to
the Faculty of the Department of Psychology
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts

by
J. Sam Quick, III
April 1971
RELATIONSHIP OF SUBJECTS' PREEXPERIMENT ATTITUDES TOWARD PSYCHOLOGY AND EXPERIMENTER PRESTIGE

APPROVED May 7, 1971:

[Signatures]

Dean of the Graduate School
Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>v</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Statement of Problem</td>
<td>12</td>
</tr>
<tr>
<td>Method</td>
<td>14</td>
</tr>
<tr>
<td>Results</td>
<td>18</td>
</tr>
<tr>
<td>Discussion</td>
<td>23</td>
</tr>
<tr>
<td>References</td>
<td>31</td>
</tr>
<tr>
<td>Appendix A</td>
<td>36</td>
</tr>
<tr>
<td>Appendix B</td>
<td>37</td>
</tr>
<tr>
<td>Appendix C</td>
<td>38</td>
</tr>
<tr>
<td>Appendix D</td>
<td>39</td>
</tr>
</tbody>
</table>
Abstract

A study is reported which examines the effects of Ss' attitudes toward psychological research as measured by the Psychology Research Survey (PRS) (high or low), apparent E prestige (high, medium, or low), and sex of S on ratings of success and failure given in Rosenthal's person-perception task. It was found that E prestige interacted with sex of S to significantly affect performance. Several of the pictures used in the person-perception task were also found to be consistently rated as depicting success or failure. Questions are raised concerning the psychometric soundness of both the PRS and the person-perception task. Implications for future research are indicated with the need for replication studies emphasized.
Acknowledgements

I would like to express my sincere appreciation to Dr. James R. Craig, my Thesis Committee Chairman, for his generous and patient giving of himself, his time, and his expertise in guiding this project. I also wish to express my sincere appreciation to Dr. Harry Robe and Mr. John Dullaert, members of my thesis committee, for their continued helpfulness and support. And a special thanks to my wife, Jackie, for her understanding and encouragement, and for her patience in typing this paper.
Introduction

In recent years behavioral scientists have become quite concerned with several previously neglected variables in psychological experiments, namely, the subjects (Ss), the experimenter (E), and the possible interactions. Considerable research has been conducted concerning the social psychology of the psychological experiment (Kessel & Barber, 1968; Masling, 1960, 1966). Sex, age, race, and a variety of E and S characteristics such as expectancies and prior experience appear to be selectively operative in many experimental conditions (Argyris, 1968; Kintz, Delprato, Mettee, Persons, & Schappe, 1965).

Subjects' Attitudes Toward Psychology. Until very recently, it was generally assumed that most human Ss reacted within the experimental environment in a relatively uniform manner, more or less like a "good" S should. In response to this oversimplified view, Orne (1962) proposed that the S's experimental behavior is a function of the total situation, which includes the experimental variables being investigated and at least one other set of variables he called demand characteristics, i.e., the idea that the S perceives certain characteristics in the experimental environment to which he attends and which demand or stipulate to him how he should respond. According to Orne, the Ss sense these demand characteristics because of their generally high regard for science
and experimentation.

Along similar lines, the work of Holmes (1967) indicated that the more introductory psychology students participated in experiments, the greater the probability that they would become aware of the reinforcement contingency in a verbal conditioning experiment. He also discovered that once they were aware of the reinforcement contingency, they would be more cooperative. Holmes additionally found, using a questionnaire, that Ss with higher levels of experimental experience perceived experiments as more scientific and valuable and professed a higher level of intended cooperation than did Ss with lower experience levels. On the basis of these findings, it was concluded that the number of previous experiments in which Ss have participated can affect their perceptions of experiments, their behavioral intentions, and their actual performance in experiments. However, Holmes stated that caution must be used when generalizing from these findings since the nature of the effect may differ from laboratory to laboratory.

Argyris (1968), taking a somewhat different perspective, contended that Ss are placed in situations similar to those that organizations create for lower level employees. By drawing on organizational theory, Argyris suggested that many "volunteer" Ss have both overt and covert hostility toward research. If the S likes the E, he cooperates; if not, the S may intentionally attempt to "botch" the experiment. Argyris considered factors such as Ss' feeling pressured to participate and a general disbelief and mistrust of
the researcher responsible for these negative attitude formations.

Adair and Fenton (1970) felt that the above stated polar positions concerning Ss' attitudes toward experimentation were too narrow. In order to more adequately account for Ss' attitudes, Adair and Fenton postulated that S differences are "due in large part" to differences in preexperiment attitudes toward psychology and psychological research which lead to a continuum of motivation to cooperate with the experimenter. Not only are Ss' attitudes toward psychology and psychological research an important variable in influencing experimental results, but they are probably functional even in those Ss who have not been previously exposed to experimentation.

Accordingly, Adair and Fenton (1970) have devised the Psychology Research Survey (PRS), which is a purported measure of preexperimental attitudes toward the discipline of psychology and psychological research. The PRS, which requires approximately 20 minutes to administer, consists of 52 items presented in a five-choice Likert format. The items are stated impersonally or in the third person so that Ss who have not previously taken part in experiments may indicate their attitudes toward psychology. Split-half reliabilities on the PRS have yielded corrected coefficients ranging from .89 to .95 (Adair & Fenton, 1970).

Research conducted thus far with the PRS has led to conflicting results. Adair and Fenton (1970) gave a group of
male and female Ss a brief and factual written report on the topic of vivisection and then measured their opinion concerning it. Next the Ss were administered the PRS and according to the distribution of scores were divided into high, low, and medium attitude groups. Upon completing the PRS the Ss were asked to listen carefully to a tape recording of a pro-vivisection message, and afterwards another measure of the Ss' opinions concerning vivisection was obtained. Finally, the Ss were asked to respond to an open-ended questionnaire designed to test their awareness of the experiment. Adair and Fenton hypothesized that Ss who had more positive attitudes toward psychology would show greater opinion change. The results indicated that Ss who were aware of the E's hypothesis and who had positive attitudes toward psychology and psychological research complied with the demand characteristics so as to confirm the E's hypothesis.

In contrast to this finding, more recent research conducted by Adair (1970) has led to essentially different results. Adair found, using a verbal conditioning procedure as the experimental task, that cooperative Ss had less positive attitudes toward psychology as measured by the PRS. Although this research has led to conflicting results, both Adair's (1970) findings and those of Adair and Fenton (1970) suggest that the S's attitude toward psychology is a significant variable in influencing the results of psychological experiments.

It appears that the PRS may prove to be an interesting
and fruitful approach to investigating S differences since Ss' PRS scores would enable their attitudes to be quantified. This would allow for a comparison of various groups of Ss and their respective attitudes toward psychology. Most importantly, it would allow for the examination and study of Ss' attitudes along the full continuum from very positive attitudinal sets to very negative ones.

In examining Ss' attitudes toward psychological research, a variable that should be considered is how the Ss are selected. A large portion of the human Ss used in psychology are drawn from the college population, a fact which raises several questions. For example, if students are required to take part in so many hours' worth of experiments, how much choice do they have as to which experiments they will participate in, and how will this affect their attitudes? Possibly one of the most important questions was raised by Rosenthal (1965): Is the volunteer S unique, and if so, what limitations are involved?

Rosenthal (1965) viewed the act of volunteering as a non-random event, determined partially by situational variables and partially by the specific personal attributes of the volunteer. On the basis of studies conducted with college students, he suggested that greater intellectual ability, interest, and motivation, greater unconventionality, lower age, less authoritarianism, and greater sociability are generally characteristic of volunteer Ss. Rosenthal further suggested that one limitation when using volunteer Ss was
that placed on subsequent statistical procedures and inference by the violation of the requirement of random sampling.

Although many investigators have reported differences between volunteers and nonvolunteers (Levitt, Lubin, & Brady, 1962; Suefeld, 1964), other investigators have found trivial or no differences (Walters & Kirk, 1969; Wilson & Patterson, 1965). For example, Walters and Kirk (1969) studied individual differences among introductory psychology students on a series of motivational-personality and biographical variables. They found no significant differences concerning this series of variables among volunteers who appeared for the experimental sessions, volunteers who did not appear for the experimental sessions, and nonvolunteers.

Influence of Experimenter on Results. While psychologists have traditionally recognized that the characteristics of an E may influence experimental behavior, it has only been in the last decade or so that the E has been seriously studied as an independent variable. Rosenthal (Barber & Silver, 1968; Rosenthal, Fode, Friedman, & Vikan-Kline, 1960), who was instrumental in bringing attention to the need of studying the E, has examined several ways in which an E might inadvertently influence the results of his research. Two possible sources of influence are the experimenter's personal attributes (age, race, status, etc.) and the experimenter's bias. Experimenter bias refers to the fact that E's may often obtain spurious results due to their expectancies, hypotheses, or desires. The present review focuses on the effect exerted
upon the experimental outcome by the personal attributes of
the E.

Kintz et al. (1965) contend that wherever an experimenter-
subject relationship exists, it is very possible for the E
to contaminate his data in one or more ways. For example,
Sanders and Cleveland (1953), using projective techniques,
reported that overtly anxious Es (as indicated by their own
Rorschach responses) tended to elicit flexibility and respon-
siveness in their Ss' Rorschach responses. Sanders and
Cleveland also found that overtly hostile Es (measured by
their own Rorschach scores) elicited more stereotyped and
passive responses and less of the hostile responses. Another
study (Brogden, 1962) reported significant differences between
Es in relation to their degree of experience and their ability
to condition Ss.

Not only the E's personality but also the S's percep-
tion of the E's personality can contribute to the experi-
menter effect. Two studies investigating Ss' perception of
the E (Rosenthal, Fode, Friedman, & Vikan-Kline, 1960;
Rosenthal & Persinger, 1962) supported the hypothesis that
naive Ss may have a predetermined "set" about what a typical
E is like--scientific, intelligent, etc.

Yagoda and Wolfson (1964) examined the degree to which
the examiner might influence the productions of a S on the
Human Figure Drawing Test. They compared the drawings of
two groups of normal female Ss. One group was examined by
males with mustaches, and the other group was examined by
males without mustaches. They found that Ss who had been examined by males with mustaches had drawn significantly more mustaches on their drawings of male figures.

As a partial solution to some of these problems, Kintz et al. (1965) suggested counterbalancing of Es and the use of factorial designs which include the E as a major independent variable. Kintz et al. also proposed "the elimination of verbal and visual cues, including inflections of the voice, speaking peculiarities, gestures, etc., as transmitted to Ss during the reading of instructions." 2317.

Not only are specific attributes of Es being considered more closely, but the actual number of Es in experiments is being given more attention. Many studies currently employ more than one E. For example, Woods (1961) determined that 48 per cent of a large sample of journal articles used two or more authors. McGuigan (1963) has noted that when more than one E has been used, techniques of control should be specified and the data analyzed and reported as a function of Es; in addition, McGuigan suggested that interactions between Es and treatments should be tested. He also stated that when only one E is used, the results cannot, strictly speaking, be generalized to a population of Es.

**Experimenter-Subject Interactions.** The controlling for individual differences among Ss and Es becomes additionally complex when various parameters of the subject-experimenter interactions are taken into consideration. Some of the most important variables in this interaction, thus far not considered in this review, are E prestige, sex, and interpersonal
attraction.

Working with an introductory psychology class, Verplanck (1956) demonstrated that when the student Es tried to reinforce plural nouns according to the Greenspoon procedure (Greenspoon, 1955), those students who were most successful were more mature, were more socially acceptable and prestigious, and made a better appearance than those Es who were unsuccessful. Marion (1956), also using a verbal conditioning procedure, demonstrated that Es with high status were more efficient reinforcers than low-status Es.

Contrary to Verplanck and Marion's observations, Ekman and Friesen (1960), using verbal conditioning of hostile responses to peer photographs as a dependent variable, found no relationship between the E's prestige as perceived by the Ss and responsivity to reinforcement. Blaufarb (1961), using a verbal conditioning task, also found no relationship between E prestige (determined by age, expertness, and educational rank of E) and the S's responsivity to reinforcement. The inconsistent results possibly indicate that certain Es are more effective reinforcers for certain Ss under certain conditions.

The influence of the E's sex and size on Draw-A-Person productions was investigated by Holtzman (1952). Two male Es were used, one of whom was nearly a foot taller and 60 pounds heavier than the other, and two female Es were used who differed in the feminine qualities they displayed. Holtzman found that none of his 12 judges could guess better
than chance either the sex or the identity of the Es by inspecting the drawings of the 40 male and 40 female Ss employed. Similar results were reported by Garfield, Bleck, and Melker (1952), who used two female and two male Es to administer the TAT to 56 female and 54 male Ss. Neither the sex of the E nor the interaction of E's and S's sex produced significant differences in the stories.

Sarason and Minard (1963), using a procedure where Ss were reinforced for emission of first person pronouns, reported that the sex of the E and the hostility (measured by a hostility scale) exhibited by the S significantly influenced the performance of Ss. Sarason and Minard also manipulated the degree of face to face contact between the S and E and the E's prestige. The E's prestige was determined by the initial method of contacting the Ss, by the experience of the E, and by the E's introduction of himself at the beginning of the actual experiment. It was found that the degree of contact between S and E and the E's prestige produced significant results. Stevenson and Allen (1964), investigating the sex variables, used a simple sorting task and found that with either male or female Es, female Ss made more responses than did male Ss. And as expected, all Ss performed better under an opposite-sexed E.

Kessel and Barber (1968), on the basis of the results of the survey of eight studies in verbal conditioning, stated that a general trend exists for positive interpersonal attraction to enhance the E's ability to influence the S through
social reinforcement and that negative interpersonal attraction tends to reduce the E's potency as a reinforcer.

Reviewing the results of studies in which interpersonal attraction was induced, Kessel and Barber (1968) also found that the induction of a positive attitudinal set in the S toward the E resulted in positive interpersonal attraction for the E. An induced positive set also resulted in increased E influence as measured by conditionability as compared with Ss in whom a negative attitudinal set was induced. In concluding their review, Kessel and Barber observed that there is no simple relationship between attraction and influence in verbal conditioning research.
Statement of Problem

From the studies reviewed thus far, it is evident that many situational and interpersonal variables affect the performance of both the S and the E in psychological research. The specific attitudes that Ss have toward psychology and psychological research prior to participation in experiments appear to be a relatively neglected area in the literature to date. The present study is an attempt to examine Ss' preexperimental attitudes toward psychology and psychological research as a possible determinant of Ss' performance in experiments. The prestige of the E and sex of the S are also examined.

All Ss were given the PRS (Adair & Fenton, 1970) prior to the actual experiment, and on the basis of their scores the Ss were divided into high and low attitude groups. One male E was used, who played three roles differing in prestige value--high, medium, and low E prestige. The prestige of the E was determined by the E's appearance and the manner in which the E initially introduced himself to the S. The sex of the S was also included as an independent variable so that possible sex differences and interactions could be investigated.

The dependent variable measure consisted of Ss' scores on the Rosenthal person-perception task (Barber & Silver, 1968), in which a series of 10 faces are rated by the S.
according to how much success or failure he feels the faces are exhibiting. This task was deemed appropriate because it has proven to be a relatively sensitive measure of various aspects of the experimental environment. Shames and Adair (1967), after studying the structure of the person-perception task, concluded that the judgment asked of the S tends to be ambiguous and that the S typically turns to the E for indication as to how he should respond.

It is expected that one's PRS score will significantly influence one's score on the person-perception task. The results are also expected to indicate that the prestige of the E will significantly affect the dependent variable. It is further hypothesized that neither the sex of the S nor the individual pictures of the person-perception task will significantly influence the results. Moreover, no significant interactions are expected.
Method

Subjects. Three Introductory Psychology classes at Western Kentucky University totaling about 200 students were administered the PRS at the beginning of the semester. On the basis of the distribution of their PRS scores, the lowest 45 and the highest 45 Ss were selected and assigned a day on which they were required to take part in the experiment. The mean PRS score for the high attitude group was 163.5 with scores ranging from 162 to 190. The mean PRS score for the low attitude group was 143.2 with scores ranging from 133 to 148. In the final analysis, data from only 60 students were used since some students neglected to participate and other data were randomly discarded in order to obtain a balanced design. Sex was balanced with 30 females and 30 males participating. Since nearly all the Ss were freshmen first being exposed to psychology, prior experimental experience was nil. It should also be recognized that the volunteer S problem was essentially eliminated since participation was required and the Ss were assigned to the particular day during which they participated. An amplification of the sample in terms of the Ss' PRS scores is shown in Appendix A.

Task. The dependent variable was the S's score on the Rosenthal person-perception task (Barber & Silver, 1966). In this task the S is asked to rate the faces of people pictured in a series of 10 relatively ambiguous black and
white photographs. The faces are rated on a scale from -10 to +10 according to whether each of the persons depicted, in the opinion of the S, has been "experiencing failure" or has been "experiencing success." A rating of -10 connotes extreme failure and a rating of +10 extreme success. The photographs have been standardized (Barber & Silver, 1968; Ducan, Rosenberg, & Finkelstein, 1969) so that in bias-free conditions the ratings tend to average 0.

**Design.** A 2x2x3x10 split-plot factorial design was employed. The between factors consisted of the S's PRS score (high or low), S's sex (male or female), and E prestige value (high, medium, or low). The within factor consisted of the 10 individual pictures of the person-perception task.

Three groups of Ss (balanced for PRS scores and sex) were randomly assigned to one of three days during which they were required to participate in the experiment. On the first of the three days the E assumed a high prestige value, on the second day he assumed a medium prestige value, and on the final day he assumed a low prestige value. The E prestige was determined by the E's appearance and by his initial introduction to the S. The exact wording of the introductions can be found in Appendix B.

1. **High prestige.** Appearance: dress shirt, conservative tie, and sport jacket; professional manner. The E introduced himself as Dr. Quick, a member of the Psychology Department, engaged with colleagues in a research project concerning various aspects of success and failure.
2. **Medium prestige.** Appearance: neat sport shirt; manner consistent with role of graduate student. The E introduced himself as Sam Quick, a graduate student currently conducting research for his Master's degree.

3. **Low prestige.** Appearance: striped bell-bottoms, t-shirt, and no shoes; casual manner, E chewing gum. The introduction of the E consisted merely of "Hi, I'm Sam."

**Procedure.** The Ss remained in a waiting room until the E individually escorted them to the experiment. After each S entered the experimental situation, the E introduced himself and played the taped instructions (female voice) for the person-perception task. The Ss were asked to rate each of the 10 person-perception faces as either experiencing failure or experiencing success on a scale from -10 to +10, -10 being extreme failure and +10 being extreme success. A copy of the instructions can be found in Appendix C. A 5x3 inch index card with a depiction of the scale was placed before the S. Each S was asked to rate one or more sample faces before the 10 faces actually comprising the dependent measure were presented. After each face was rated, the E responded with an "OK" regardless of the rating. The Ss verbalized their ratings, and the E recorded them. After the administration of the task, each S was debriefed according to the format presented in Appendix D. Essentially, the debriefing consisted of assuring that the S had received no relevant information about the experiment prior to his taking part and of asking his cooperation in not mentioning.
anything about the experiment for a period of several weeks.

**Scoring and Analysis.** The ratings on the person-perception task were recorded at the time of the experiment as both positive and negative integers; in other words, the -10 and +10 scale was used. However, in order to expedite the statistical procedures, the scores were transformed to a positive integer scale with \(-10 = 1, -9 = 2, \) etc. so that the revised scale varied from 1 to 21.

A 2x2x3x10 factorial analysis of variance was used to analyze the results of the experiment. Where multiple comparisons were conducted, Duncan's Multiple Range Test (Edwards, 1968) was employed.
Results

The analysis indicated that E's prestige interacted significantly with the sex of the S, \( F(2,48) = 4.47, p<.05 \), as shown in Figure 1. The analysis is summarized in Table 1. Examination of the interaction utilizing Duncan's Multiple Range Test (Edwards, 1968) indicated that the high and medium E prestige-female Ss rated the pictures significantly higher (\( p<.05 \)) than the low E prestige-female Ss. The male Ss' ratings were intermediate between the high and medium E prestige-female Ss and the low E prestige-female Ss and did not differ significantly from either extreme. The means and the summary of the multiple comparison are presented in Table 2.

The individual pictures were found to vary significantly among themselves, \( F(9,432) = 35.60, p<.01 \). Using the Duncan multiple comparison procedure, Pictures 4 and 10 were found to be rated significantly higher than the other pictures (\( p<.05 \)). On the other hand, Picture 9 was found to be consistently rated lower than any of the other pictures. The means of the other pictures were found to exist in overlapping groupings of non-significantly different means: the means for the groups of Pictures 2, 5, and 1; 5, 1, 6, and 8; 6, 8, 3, and 7 were found to be not significantly different. The Duncan's analysis is summarized and presented in Table 3.
Fig. 1. Experimenter Prestige-Subject's Sex Interaction.
TABLE 1
Analysis of Variance
(N=60)

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>E Prestige (A)</td>
<td>129.0</td>
<td>2</td>
<td>64.5</td>
<td>2.26</td>
</tr>
<tr>
<td>S's PRS Score (B)</td>
<td>.7</td>
<td>1</td>
<td>.7</td>
<td>&lt;1</td>
</tr>
<tr>
<td>S's Sex (C)</td>
<td>20.9</td>
<td>1</td>
<td>20.9</td>
<td>&lt;1</td>
</tr>
<tr>
<td>A x B</td>
<td>30.2</td>
<td>2</td>
<td>15.1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>A x C</td>
<td>254.9</td>
<td>2</td>
<td>127.5</td>
<td>4.47*</td>
</tr>
<tr>
<td>B x C</td>
<td>22.4</td>
<td>1</td>
<td>22.4</td>
<td>&lt;1</td>
</tr>
<tr>
<td>A x B x C</td>
<td>23.3</td>
<td>2</td>
<td>11.7</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Ss within Groups</td>
<td>1369.3</td>
<td>43</td>
<td>31.8</td>
<td></td>
</tr>
<tr>
<td>Pictures (D)</td>
<td>4645.6</td>
<td>9</td>
<td>516.2</td>
<td>35.60**</td>
</tr>
<tr>
<td>A x D</td>
<td>194.0</td>
<td>18</td>
<td>10.8</td>
<td>&lt;1</td>
</tr>
<tr>
<td>B x D</td>
<td>233.2</td>
<td>9</td>
<td>26.5</td>
<td>1.83</td>
</tr>
<tr>
<td>C x D</td>
<td>213.8</td>
<td>9</td>
<td>23.8</td>
<td>1.64</td>
</tr>
<tr>
<td>A x B x D</td>
<td>202.6</td>
<td>18</td>
<td>11.3</td>
<td>&lt;1</td>
</tr>
<tr>
<td>A x C x D</td>
<td>308.5</td>
<td>18</td>
<td>17.1</td>
<td>1.18</td>
</tr>
<tr>
<td>B x C x D</td>
<td>124.7</td>
<td>9</td>
<td>13.9</td>
<td>&lt;1</td>
</tr>
<tr>
<td>A x B x C x D</td>
<td>406.3</td>
<td>18</td>
<td>22.6</td>
<td>1.56</td>
</tr>
<tr>
<td>D x Ss within Groups</td>
<td>6259.9</td>
<td>432</td>
<td>14.5</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>14,441.3</td>
<td>599</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.--*p<.05, **p<.01.
### TABLE 2

Duncan's Multiple Range Test
for \( E \) Prestige-S's Sex Interaction
(N=60)

<table>
<thead>
<tr>
<th></th>
<th>L/F</th>
<th>M/M</th>
<th>H/M</th>
<th>L/M</th>
<th>M/F</th>
<th>H/F</th>
<th>Shortest Sig. Ranges (.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means</td>
<td>10.02</td>
<td>10.89</td>
<td>11.22</td>
<td>11.49</td>
<td>12.21</td>
<td>12.49</td>
<td></td>
</tr>
<tr>
<td>L/F</td>
<td>10.02</td>
<td></td>
<td>1.20</td>
<td>1.47</td>
<td>2.19*</td>
<td>2.47*</td>
<td>R2 1.50</td>
</tr>
<tr>
<td>H/M</td>
<td>11.22</td>
<td></td>
<td>.33</td>
<td>.60</td>
<td>1.32</td>
<td>1.60</td>
<td>R3 1.57</td>
</tr>
<tr>
<td>M/M</td>
<td>10.89</td>
<td></td>
<td>.33</td>
<td>.60</td>
<td>1.32</td>
<td>1.60</td>
<td>R4 1.63</td>
</tr>
<tr>
<td>L/M</td>
<td>11.49</td>
<td></td>
<td>.27</td>
<td>.99</td>
<td>1.27</td>
<td></td>
<td>R5 1.66</td>
</tr>
<tr>
<td>M/F</td>
<td>12.21</td>
<td></td>
<td></td>
<td>1.72</td>
<td>1.00</td>
<td></td>
<td>R6 1.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.28</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. -- The following designations are used: High \( E \) prestige-Male (H/M), High \( E \) prestige-Female (H/F), Medium \( E \) prestige-Male (M/M), Medium \( E \) prestige-Female (M/F), Low \( E \) prestige-Male (L/M), and Low \( E \) prestige-Female (L/F).

Note. -- Any two means underscored by the same line do not differ significantly. Any two means not underscored by the same line do differ significantly.
TABLE 3

Duncan's Multiple Range Test for the 10 Individual Pictures of the Person Perception Task

<table>
<thead>
<tr>
<th>PICTURES</th>
<th>2</th>
<th>5</th>
<th>1</th>
<th>6</th>
<th>8</th>
<th>3</th>
<th>7</th>
<th>10</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means</td>
<td>8.87</td>
<td>9.45</td>
<td>9.55</td>
<td>10.98</td>
<td>11.07</td>
<td>12.02</td>
<td>12.50</td>
<td>15.92</td>
<td>16.30</td>
</tr>
<tr>
<td>9</td>
<td>7.22</td>
<td>1.65*</td>
<td>2.33*</td>
<td>2.33*</td>
<td>3.76*</td>
<td>3.85*</td>
<td>4.80*</td>
<td>5.28*</td>
<td>8.70*</td>
</tr>
<tr>
<td>2</td>
<td>3.87</td>
<td>.58</td>
<td>.68</td>
<td>2.11*</td>
<td>2.20*</td>
<td>3.15*</td>
<td>3.63*</td>
<td>7.05*</td>
<td>7.43*</td>
</tr>
<tr>
<td>5</td>
<td>9.45</td>
<td>1.00</td>
<td>1.53</td>
<td>1.62</td>
<td>2.57*</td>
<td>3.05*</td>
<td>6.17*</td>
<td>6.95*</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>9.55</td>
<td>1.43</td>
<td>1.52</td>
<td>2.47*</td>
<td>2.95*</td>
<td>6.37*</td>
<td>6.75*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>10.98</td>
<td>.09</td>
<td>1.04</td>
<td>1.52</td>
<td>4.94*</td>
<td>5.32*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>11.07</td>
<td>.95</td>
<td>1.43</td>
<td>4.85*</td>
<td>5.23*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>12.02</td>
<td>.43</td>
<td>2.20*</td>
<td>4.28*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>12.50</td>
<td>3.42*</td>
<td>3.80*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>15.92</td>
<td>.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.—Any two means underscored by the same line do not differ significantly. Any two means not underscored by the same line do differ significantly.
Discussion

Effect of Subject's Attitude Toward Psychology As Measured by PRS. The finding that one's PRS score did not significantly affect one's person-perception ratings is somewhat at odds with prior research. Adair and Fenton (1970) found that Ss who had positive attitudes toward psychological research (high PRS scores) complied with the demand characteristics of the experiment. On the other hand, Adair (1970) found essentially the opposite, namely, that the cooperative Ss exhibited less positive attitudes toward psychology.

Adair and Fenton (1970) proposed that the construct validity of the PRS scale would be determined by demonstrating a relationship between scaled scores and external criteria. The present study, while not necessarily detracting from the construct validity of the PRS, certainly lends no support in establishing this validity.

There is one consideration that should be taken into account when comparing the present study with prior research using the PRS. Adair (1970) referred to mean PRS scores for one group of 196 and for another group of 207.92. The low and high group means for the present study (Appendix A) were 143.2 and 168.5. In effect, then, the high PRS mean in the present study is 27.5 points lower than the low PRS group mean in Adair's study. This suggests that there is a lack of consistency in operationally defining positive and negative
attitudes toward psychology and psychological research based on PRS scores. Adair's Ss were also drawn from introductory courses; however, whether his Ss were exposed to any experimentation previous to his study is not known. Perhaps sampling differences or regional geographical differences in Ss may help to explain this situation.

Based on the studies thus far conducted using Adair and Fenton's PRS, it seems desirable that the PRS be given a thorough psychometric examination before additional research using this instrument is undertaken. In particular, the PRS should be standardized in terms of what scores constitute positive and negative attitudinal sets. It would also be worthwhile to determine if the PRS is, as Adair and Fenton (1970) have suggested, a valid measure of the attitudes of Ss who have no prior experimental experience.

While discussing Ss' attitudes toward psychology and psychological research, a few related matters deserve attention. Investigators should consider using some method of follow-up concerning the Ss' feelings regarding the experiment subsequent to their taking part. For example, what demand characteristics, if any, do Ss sense prior to or during the experiment? Such post-experimental information could be obtained by administration of a questionnaire immediately after the S participated.

The amount of valuable data that could be obtained from more adequately sampling Ss' attitudes toward experimentation should not be overlooked. It is very possible that Ss have
much to teach researchers. For instance, it has been suggested (Argyris, 1968) that some Ss have negative attitudes toward psychological research. Given that this is the case, it seems that it would be helpful to discover why some students have these negative attitudes. Did they learn them in the classroom, or were their prior experimental experiences unpleasant? If prior experimental experiences did leave the student with a negative attitudinal set, why was this the case?

Effect of Experimenter Prestige. The somewhat surprising finding that the prestige of the E did not influence the ratings on the relatively sensitive person-perception task is in contrast with the findings of Sarason and Minard (1963) and Verplanck (1965). However, Ekman and Friesen (1960) reported similar results to those of the present study, stating the E's prestige as perceived by the Ss did not influence the Ss' responsivity to verbal reinforcement.

An undoubtedly fruitful approach in eliminating some of the existing confusion would be to engage in more replication studies. For example, if the present study were completely replicated, so that the only variable that would change would be the E himself, an estimate of the reliability of the present findings would be possible. For instance, if different results were obtained on several replications, it could be stated with some degree of certitude that the E appeared to be the crucial variable, other conditions remaining constant. Replication of various findings would insure more defensible
and reliable data.

When E prestige is treated as an independent variable in future research, an attempt might be made to distinguish between the operational definition of E prestige and E prestige as perceived by the S. These two values of prestige could then be correlated. It might also be helpful to videotape the experimental sessions. This would allow for the detection of unintentional paralinguistic or kinesic cues operating during the experiment.

Future studies concerning the experimenter-subject interaction should also take into consideration the following points. Using one E is functional both for practical reasons and as a method of control. It should be recognized, however, that when only one E is employed, the results, strictly speaking, cannot be generalized beyond the E used in the study (McGuigan, 1963). Although the experimental design would become more complicated, using several Es could enhance the possibility of valid generalizations and increase our knowledge of the experimenter effect as well.

Effect of Subject's Sex. Sex as a main effect was non-operative, but there was an interaction between the low E prestige-females and low E prestige-males. Apparently, females reacted most negatively to the low E prestige condition, whereas males reacted about the same regardless of E prestige. As far as the medium and the high E prestige conditions were concerned, the present results are consistent with the findings of Stevenson and Allen (1964); namely,
females perform more positively under an opposite-sexed E.

Despite the abundance of research in most areas concerning the sex variables, more investigation of the sex variables in the area of experimenter-subject interaction is needed. For example, in the present study it was found that the sex of the S interacted with E prestige. The assumption is then implicitly made after examining the interaction that the female Ss reacted negatively to the low E prestige condition. However, it may have been that the female Ss were actually reacting to the appearance of the E in a t-shirt, rather than to the low E prestige as operationally defined. In other words, were the Ss reacting to what was considered the pertinent dimension—low E prestige, or were their reactions determined by something peculiar to the E in the low prestige role? A follow-up questionnaire could be helpful in the investigation of such issues.

Once again, replication studies are needed to establish the reliability of past findings. As an example, investigators need reliable data determining which types of experimental problems are susceptible to confounding of the sex variables and which are not. With information of this nature, it would then be possible to tackle such problems as why different tasks are differentially influenced by sex variables.

**Effect of Individual Person-Perception Pictures.** Perhaps one of the most significant findings of the present research is the indication that several of the faces that comprise the Rosenthal person-perception task appear to possess a bias
rather than being bias-free, as is generally assumed (Duean, Rosenberg, & Finkelstein, 1969). It would be interesting to see if the biases detected in the present study would be similar in other experimental situations. In other words, are these biases reliable? Barber and Silver (1968) assert that determining the reliability of the person-perception task is especially important since reliability may be regarded as an independent variable. For example, it is important to know whether the experimenter effect operates on reliable measures as well as on unreliable ones.

Dana and Dana (1969) state that the person-perception task was constructed by removing those 20 faces most often rated zero from the original 57 stimuli. Since this original construction the faces have not been administered in other contexts to determine whether or not they would still be rated on the average as zero. A restandardization of the person-perception task seems in order.

Although a great deal of the recent research done in the E effect area has used the Rosenthal person-perception task, many other tasks have been used. Rosenthal (1969), after reviewing more than 40 studies, reports that tasks other than the person-perception task tend to show greater E effects. Possibly one of the many tasks Rosenthal mentions, such as reaction time or tone length discrimination, might prove to be more suited to future research concerning E effects than the person-perception task would be. However, the elimination of the biased faces (if, indeed, future
research confirms the possibility of biases) would probably suffice in justifying the continued use of Rosenthal's task.

An interesting and recent alternative to the person-perception task is suggested by Bootzin (1970). He proposes a cue utilization paradigm in which a detailed analysis of the S's performance is possible. Bootzin's model represents the judgment process as one in which a criterion is inferred from several cues which are related to the criterion in a probabilistic fashion. Within this framework the cues which the S relies upon can be specified, and it is also possible to analyze the S's cue dependencies before, during, and after influence attempts. Such a paradigm offers a greater range of potential than the person-perception task, which provides for only the most global analyses.

The value of a paradigm like Bootzin's cue utilization can better be realized if a trend noted in the present study is mentioned. Several of the Ss, as they rated the individual faces in the person-perception task, made comments like "He seems to be happy (or unhappy) so I guess he's pretty successful (unsuccessful)." Obviously, such inferential logic could easily be tainted by the S's own feelings of happiness (a volatile concept) at the time of the experiment. In other words, what exactly does the person-perception task measure? And does the measure consistently tap the same thing? It is time that the reliability and validity of the often used person-perception task be determined.

**Summary.** The ignoring of situational variables and
individual differences associated with Ss and Es can be hazardous research methodology. Implications are fairly straightforward since the information available to date permits few clear-cut conclusions. Rigorous research is needed which varies the characteristics of the participating individuals (Es and Ss), the relationship between these individuals, and the types of experimental tasks used.

Many of the various points previously touched upon emit a recurrent theme—a theme resounded by several researchers including Barber and Silver (1968):

The behavioral sciences should adopt a rule that has long been prevalent in the physical and biological sciences, namely, research results should not be accepted until they have been replicated in independent laboratories [p. 27].
References


Orne, M. T. On the social psychology of the psychology experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 1962, 17, 776-783.


Rosenthal, R., Fode, K. L., Friedman, C. J., & Vikan-Kline, L. L. Subjects' perception of their experimenter under


Verplanck, W. S. The operant conditioning of human motor


Wilson, P. R., & Patterson, V. Sex differences in volunteering behavior. Psychological Reports, 1965, 16, 976.


Appendix A

Amplification of Sample in Terms of Subjects' PRS Scores (N=60)

<table>
<thead>
<tr>
<th></th>
<th>High PRS Group</th>
<th>Low PRS Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>Overall</td>
<td>168.5</td>
<td>143.2</td>
</tr>
<tr>
<td>Males</td>
<td>165.6</td>
<td>142.9</td>
</tr>
<tr>
<td>Females</td>
<td>171.4</td>
<td>143.5</td>
</tr>
</tbody>
</table>

Range:
- High PRS Group: 162-190
- Low PRS Group: 133-148
Appendix B

Experimenter's Introduction for Each Level of Experimenter Prestige

High Experimenter Prestige

Hello. Have a seat please. I'm Dr. Quick. I'm doing some research along with several other members of the Psychology Department concerning various aspects of success and failure. This particular experiment is one of a series and should only involve about five minutes. Your name is (Miss/Mr.) ________________________ and you're in ________________________ class. OK, fine. Please listen carefully to the following instructions. [The instructions were played.]

Do you understand? Good. For example, how would you rate this face [practice face]? OK. Now give me a rating on each of the following.

Medium Experimenter Prestige

Hello. Have a seat please. My name is Sam Quick. I'm currently conducting research for my Master's degree. This experiment will only take about five minutes. Please listen carefully to the following instructions. [The instructions were played.]

Do you understand? Good. For example, how would you rate this face [practice face]? OK. Now give me a rating on each of the following.

Low Experimenter Prestige

Hi, I'm Sam. This won't take long. Listen to these instructions please. [The instructions were played.]

Do you understand? Good. For example, how would you rate this face [practice face]? OK. Now give me a rating on each of the following.
Instructions Presented to Each Subject for the Person-Perception Task

"You'll be shown a series of photographed faces. After you look at each photograph, rate the face as either experiencing failure or as experiencing success on a scale from -10 to +10, -10 being extreme failure and +10 being extreme success. Look at the sample rating scale in front of you. Let's try one example."
Appendix D

Subject Debriefing Format

"Did you hear anything pertaining to this experiment before taking part today? If so, find out what and write it down."

"This completes the major portion of the study. I want to ask your help in finishing this research. It is crucial that students do not know exactly what takes place during the experiment prior to their taking part. So I want to ask you to help me by not mentioning anything about the experiment for, say, three weeks. By that time, I'll be done and you may tell anyone about what transpired here. In the meantime if anyone does ask about the experiment, please say something vague like 'A man asked me to rate some pictures,' but please don't get anymore specific than that. OK? Get a verbal commitment."