Summer 1974

THE GOOD VERSUS THE APPREHENSIVE SUBJECT: THE EFFECTS OF HYPOTHESIS AWARENESS AND EVALUATION APPREHENSION ON COMPLIANCE WITH MANIPULATED DEMAND CHARACTERISTICS OF A CONDITIONING EXPERIMENT

DEBORAH HELEN DU NANN

Follow this and additional works at: https://scholars.unh.edu/dissertation

Recommended Citation
https://scholars.unh.edu/dissertation/1062

This Dissertation is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact nicole.hentz@unh.edu.
THE GOOD VERSUS THE APPREHENSIVE SUBJECT: THE EFFECTS OF HYPOTHESIS AWARENESS AND EVALUATION APPREHENSION ON COMPLIANCE WITH MANIPULATED DEMAND CHARACTERISTICS OF A CONDITIONING EXPERIMENT

Keywords
Psychology, Experimental
INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in “sectioning” the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
DU NANN, Deborah Helen, 1947-
THE GOOD VERSUS THE APPREHENSIVE SUBJECT:
THE EFFECTS OF HYPOTHESIS AWARENESS AND
EVALUATION APPREHENSION ON COMPLIANCE
WITH MANIPULATED DEMAND CHARACTERISTICS
OF A CONDITIONING EXPERIMENT.

University of New Hampshire, Ph.D., 1974
Psychology, experimental

Xerox University Microfilms, Ann Arbor, Michigan 48106
THE GOOD VERSUS THE APPREHENSIVE SUBJECT: THE EFFECTS OF HYPOTHESIS AWARENESS AND EVALUATION APPREHENSION ON COMPLIANCE WITH MANIPULATED DEMAND CHARACTERISTICS OF A CONDITIONING EXPERIMENT

by

DEBORAH H. DU NANN
B.A., Grinnell College, 1969
M.A., University of New Hampshire, 1972

A THESIS

Submitted to the University of New Hampshire
In Partial Fulfillment of
The Requirements for the Degree of

Doctor of Philosophy
Graduate School
Department of Psychology
August, 1974
This thesis has been examined and approved.

Thesis director, Stephen J. Weber, Asst. Prof. of Psychology

Loren Cobb, Asst. Prof. of Sociology

G. Alfred Forsyth, Asso. Prof. of Psychology

Leslie A. Fox, Asst. Prof. of Psychology

Raymond L. Erickson, Dean of the Graduate School

Daniel C. Williams, Asst. Prof. of Psychology

July 19, 1974
Date
ACKNOWLEDGEMENTS

The author wishes to express her deepest appreciation to the following people for their energies in enabling the dissertation to be completed:

The members of her committee: Drs. Loren Cobb, G. Alfred Forsyth, Leslie A. Fox, Raymond L. Erickson, Stephen J. Weber, and Daniel C. Williams, who served at various times of night and day as invaluable sources of know-how, perspective, and advice. Of these, she would like to single out Stephen J. Weber, who as chairman and close friend was encouraging, critical, insightful, and compassionate for more than two years of consultation and partnership, and G. Alfred Forsyth, who in his consistently conscientious fashion, gave much of himself in study and suggestions,

Jim Cerney, for his wit and wisdom as consultant for the computer analyses,

John Rosborough, for risking moral wrath by serving as the accomplice in the experiment,

Dan Swift, for his help in making the tape recordings used in the experiment, and for supplying chocolate frappes in the hour of need,

Roger Adams, for his help in scoring the measures,

Stephanie Ackerman, who typed 'till the birds sang on more than one occasion while miraculously retaining
her refreshing sense of humor,

    Randy Easton, Gayle Scroggs, and Dick Kushner, three close friends who willingly and patiently endured throughout the endeavor by listening to scores of half-baked ideas, complaints, jargon, and other inevitable by-products of dissertation writing,

    and countless other unnamed persons who befriended and inspired her throughout the doctoral program at the University of New Hampshire.

    This dissertation is fondly dedicated to my parents, Mr. and Mrs. Roy P. Du Nann, whose sacrifices throughout my schooling have ultimately made this thesis possible.
TABLE OF CONTENTS

LIST OF TABLES .................................................. vii
LIST OF FIGURES .................................................. viii
ABSTRACT ............................................................ ix

I. INTRODUCTION .................................................. 1
   1. The Human Subject—Subject of Controversy .................. 5
   2. Recent Formulations of Subject Reactions ................... 7
       Empirical Evidence for the Good Subject Role .......... 10
       Other Roles ................................................. 16
   3. Evaluation Apprehension and the Defensive Subject ....... 20
   4. Evaluation Apprehension versus the Good Subject ......... 29
   5. The Classical Conditioning of Attitudes Paradigm .......... 44
   6. The Study of Post-Experimental Awareness .................. 50
   7. Overview of Research Objectives .............................. 59

II. METHOD .......................................................... 63
   1. Design ....................................................... 63
       Factor 1: Contingency Information (CI) ................. 63
       Factor 2: Test (T) .............................................. 64
       Factor 3: Accomplice (A) ................................. 64
       Factor 4: Social Desirability (SD) ....................... 64
   2. Subjects ....................................................... 64
   3. Procedure ...................................................... 65
**LIST OF TABLES**

1. Effect of Evaluation Cues on Dulaney-Type Multiple Choice Item of Evaluation Awareness .... 85
2. Effect of Evaluation Cues on Essay Item of Evaluation Awareness .................................. 86
3. Effect of Evaluation Cues on Attributed Gullibility ................................................. 88
4. Effect of Evaluation Cues on Attributed Conformity .................................................. 89
5. Effect of Contingency Information on Dulaney-Type Item of Evaluation Awareness ............... 91
6. Effect of Contingency Information on Attributed Gullibility and Conformity ..................... 92
7. Obtained Means for Conditioning with Ns Indicated .................................................... 96
8. Summary Table for Four Factor Analysis of Variance on Conditioning ............................... 98
9. Simple Effects of Test x Accomplice x Information x Social Desirability Interaction on Conditioning ................................................................. 99
10. Intercorrelations Among Awareness Measures at Selected Levels of Hypothesis Awareness .......... 104
11. Intercorrelations Among Awareness Measures ...... 106
12. Correlations Between SD and EA at Separate Evaluation Cue and Information Conditions ...... 108
13. Distributions of Contingency and Hypothesis Awareness ............................................. 124
LIST OF FIGURES

1. Pictorial Representation of Research Objectives ........................................ 62
2. Distribution of SD Scores with Median Split Indicated .................................. 72
3. Predicted Interaction of Independent Variables on Conditioning Based on Pilot Work .......... 78
4. Refined Predicted Interaction of Independent Variables on Conditioning Based on Manipulation Checks .......................................................... 95
5. Obtained Means for Conditioning with Significant Simple Effects Indicated .................. 97
6. Correlations Between Model Components when HA x MC Partialled Out .................. 109
7. Correlations Between Model Components when HA Scored as a Dichotomy and Resulting HA x MC Term Partialled Out ........................................ 126
ABSTRACT

THE GOOD VERSUS THE APPREHENSIVE SUBJECT: THE EFFECTS OF HYPOTHESIS AWARENESS AND EVALUATION APPREHENSION ON COMPLIANCE WITH MANIPULATED DEMAND CHARACTERISTICS OF A CONDITIONING EXPERIMENT

by

DEBORAH H. DU NANN

The dissertation reports research designed to test the relative importance of two subject motivations in a typical laboratory experiment. To investigate whether subjects are more concerned about helping the experimenter verify his predictions or presenting themselves positively, several procedural manipulations were made which affected subjects' hypothesis and evaluation awareness of the experiment.

In order to investigate the role of cooperation with these demand characteristics, several relationships between data and variables were examined. The first involved the effect of these manipulations on post-experimental reports of awareness, the second involved the effects of these manipulations on conditioning, and the third involved the relationship between awareness reports and
conditioning. In this manner, a triangulation of observations was made to provide an extensive picture of how a subject responds in the classical conditioning of attitudes experiment, and what sorts of processes underlie his/her cooperation.

Data concerning the effects of the cue manipulations on awareness indicated both convergent and discriminant validity of the hypothesis and evaluation cues. The resulting conditioning showed that both types of subject roles depend upon a complex interaction of the various cues available in the situation and individual characteristics of the subject. Both cooperation and apprehension was observed, and subjects' need for social approval enhanced the effectiveness of the evaluation cues while it decreased the effectiveness of the hypothesis cue. Post-experimental awareness measures indicated that hypothesis awareness was essential in mediating the effects of the evaluation cues, but that predictions derived from a model of conscious propositional control were not verified.

Recommendations from the study point to the need for experimental psychologists working with human subjects to consider awareness reports when interpreting their studies, to minimize the possibility of cues which produce hypothesis and evaluation awareness, to concern themselves less with the relative importance of subject "roles" and to begin specifying precise situational events which elicit or suppress compliance, and thereby extend the present research to other experimental paradigms.
I. INTRODUCTION

Although experimental psychology is marked by considerable diversity of assumptions among its practitioners, there are two basic assumptions which, while usually not verbalized, are critical to its scientific status. They are 1) that variables which are manipulated in the laboratory are specifiable, and 2) that the phenomena observed there can be generalized in some way to non-laboratory conditions. These are essential assumptions, without which the science of behavior would be useless, if not impossible. The laboratory, then, has often been assumed to be a neutral surrounding, enabling the manipulation and control of specified stimuli, as well as the prediction and measurement of resulting behavior.

There has been in recent years, however, a growing doubt as to the general legitimacy of this model for the study of human behavior. Unlike the experiments in classical physics, after which experimental psychology has self-consciously modeled itself, the central object of psychological study, the behaving organism, cannot be placed in a vacuum condition which stands unaffected by the particular experimental situation. This notion has been expressed for many years (Koch, 1959; Rosensweig, 1933) but a recent surge of empirical activity has emerged to make the social psychology of the experiment an important concern of the 1970's
(Adair, 1973; Friedman, 1967; Miller, 1972; Rosenthal & Rosnow, 1969). In this manner, experimental psychology is beginning to resemble more contemporary physics in which random variables have come to be considered as fixed variables, and psychologists are expanding their conception of relevant predictors of behavior, namely laboratory effects.

Several distinct areas of research have been born in this regard. For example, Robert Rosenthal (1966) has amassed substantial evidence that the experimenter exerts powerful effects, through subtle means, on the subjects' behavior by merely possessing an expectation about the effect an experimental condition should have. While some of his research has received cogent methodological criticism (Barber & Silver, 1968), it is generally acknowledged (Adair, 1973; Fishbein & Ajzen, 1972; Miller, 1972; Kessel, 1971) that the experimenter can be a potent source of artifact, and research into the effects of experimenter expectations continues to flourish (for example, Adair, 1973; Bloom & Tesser, 1971; Duncan, Rosenberg, & Finkelstein, 1969; Friedman, 1967; Harris, 1971; Rosenberg, 1973; Sattler, 1970).

Other programs of research have centered on the special nature of volunteer subjects (Rosenthal & Rosnow, 1969; Rosnow & Rosenthal, 1970) on the methodological confounds produced by deception (Brock & Becker, 1966; Cook, Bean, Calder, Frey, & Krovetz, 1970) on special cues in the experiment which tell subjects what is expected of them (Orne, 1962; 1968; 1969) and on laboratory-produced motiva-
tions and defenses on the part of subjects (Rosenberg, 1965, 1969). Reviews of research on these topics have recently been collected in a volume by Rosenthal and Rosnow (eds.) entitled *Artifact in Behavioral Research*.

The title of this book indicates the current status of these social psychological events in the laboratory: they are considered dangerous confounding variables which infect the true phenomena under study. In this vein, Silverman and Schulman (1970), for example, have outlined some of their effects in laboratory studies of attitude change and have made several suggestions for future research on the basis of their findings. The first is that experimenters should make better attempts to disguise the true purpose of their studies by increasing the quality and quantity of deception procedures. The second is that more studies should be conducted outside the laboratory, in non-reactive settings (Webb, Campbell, Schwartz, & Sechrest, 1966). These two procedures would apparently reduce the likelihood that the subject would respond to the specific social stimuli of the laboratory rather than to the inferred independent variable.

While these recommendations seem reasonable and would certainly solve some of the problems described, strict adherence to these solutions would leave most of the issues far from solved. First of all, changing the nature and increasing the intensity of deception procedures raises serious ethical issues (Baumrind, 1964; Kelman, 1967) as well as substantial methodological problems, as demonstrated in
the recent work on deception and suspiciousness (Allen, 1966; Cook et al., 1970; McGuire, 1969; Stricker, Messick & Jackson, 1969). It is likely that increased deception creates heightened suspiciousness, which then necessitates more deception, etc., eventuating in an endless spiral of lies and deceit, a polluted subject pool, and biased data. While the second recommendation, that of field studies, would certainly bypass some of these problems, such settings are frequently impossible for well-controlled investigations. Thus, there seems little way of avoiding the social nature of psychological research.

A more promising approach would seem to be to attack these problems directly by attempting to discover various situational antecedents in the laboratory which give rise to these artifacts. In this context, the artifact then becomes a legitimate topic of study in its own right. Noting that it is the wise researcher who realizes that at any given time one man's artifact may be another man's main effect, McGuire (1969) has traced the history of artifacts through three stages. They are 1) ignorance, when the artifact contaminates the research without the experimenter's knowledge; 2) coping, during which the experimenter becomes over-concerned and diverted in trying to control for it and eliminate it from his studies; and 3) exploitation, when the researcher recognizes the artifact as an interesting phenomenon in itself and designs studies to investigate the underlying processes. Thus, while much recent research has been primarily of stage
2, it seems time to move to stage 3.

This would be particularly the case with phenomena which we would expect to have generalizability to non-laboratory conditions. That is, artifacts generated by strictly-laboratory stimuli, such as special equipment, would be much less interesting than artifacts generated by social-psychological phenomena, like the subject's response to perceived expectancies. In this respect, the study of the experiment from the point of view of the subject is likely to result in the understanding of social processes which go far beyond the laboratory. To be sure, the subject is in a special situation—a experiment—but he brings to the experiment motives and attitudes which he is likely to bring to other social-contract situations that have properties in common with the psychological experiment.

Actually, the human subject is really the chief concern of all artifact study, for all of these artifacts bias data by affecting subjects. In order to study any social psychological aspects of the experiment, we must know more about how subjects perceive it, what sorts of cues they use to make inferences about it, and what sorts of conditions cause the various subject motivations.

The Human Subject—Subject of Controversy

This concern with the subject and his experience in the laboratory represents an interesting synthesis of historic themes in experimental psychology. As Schultz (1969) has noted, since the days when the Functionalist and Behav-
ioristic movements in the United States helped demote the human subject from the lofty position he enjoyed in Wundt's and Titchener's laboratories, the human subject has been seen as a passive input-output machine. His role was to passively and naively be subjected to stimuli and to reflexively respond to them (Boring, 1953). His very name, "subject," reflects this designation, and it differs significantly from Wundt's "reagent" who took a more active and sophisticated role in describing the effects of stimuli upon his consciousness.

The attention to social aspects of the laboratory, however, points to the need for us to consider the subject as something more than a naive and passive body. Now he is being conceptualized as actively engaged in a problem-solving task when he enters the experiment. He is not merely subjected to stimuli, but he is consciously trying to discern what is going on in the experiment, what the experimenter is trying to demonstrate, what the expected behavior means, and whether he should respond accordingly.

While some aspects of this position may have been overstated to represent a currently-fashionable humanistic concern with the importance of the subject, the recent attention to the phenomenological viewpoint of the subject represents an interesting communion of two formerly opposing camps in contemporary psychology (Schultz, 1969). On the one hand, the strict experimentalist has traditionally been rooted in the behavioristic tradition, and has thus been more
concerned with overt behavior than mediating mental processes. He has been in sharp contrast with the more recently emerging humanistic camp of psychologists who place great importance on the person's phenomenological perspective, and insist that analysis of overt behavior is at best incomplete without an exploration into the human being's phenomenological perspective, his awareness, feelings, ideas, goals, etc. (Jourard, 1967; 1968). Traditionally, these approaches to the study of human behavior have been divided along experimental vs. clinical boundaries, but now it seems possible to unite both the experimentalists' need for objectivity with the phenomenologists' regard for immediate experience, through the laboratory study of behavior and awareness. In this manner, experimental psychology would seem to be evolving in a spiral (Kuhn, 1962) with regard to the older Functionalist-Structuralist debate (Baldwin, 1895) as to who our subjects should be and how they should be studied. That is, rather than swinging back to the earlier opposite of behaviorism (that of using trained introspectionsists as subjects) we are moving forward to the inclusion of both phenomenological reports and overt behavior as converging sources of data.

**Recent Formulations of Subject Reactions**

Subject motivations have been long recognized as potential sources of bias (Rosensweig, 1933; Titchener, 1895). In non-laboratory settings, for example, the Hawthorne studies demonstrated that just being in an experiment pro-
duces special motivation on the part of workers which make it difficult to surmise inferences about the effects of environmental stimuli such as lightning since any environmental change produced better performance. Presumably the special attention was more important than the independent variables which the psychologists were studying (Roethlisberger & Dickson, 1937).

In large part, however, recent attention to the social aspects of the laboratory experiment was stimulated by Martin Orne's (1962) suggestion that human subjects come into laboratories seeking to discover what it is that is expected of them. Actively trying to discern the experimenter's expectations, subjects search for and respond to demand characteristics, which Orne defined as "the totality of cues which convey an experimental hypothesis." Broadly conceptualized, these cues were said to include the rumors or campus scuttlebutt about the research, the information conveyed during the original solicitation, the person of the experimenter, and the setting of the laboratory, as well as all explicit and implicit communications during the experiment proper [p. 779].

According to Orne, what makes these demand characteristics so pernicious to valid psychological research is that subjects are by nature motivated to help the scientist with his work in furthering scientific knowledge. That is, subjects want to be "good" subjects: they come into the laboratory concerned about having their data be valid and useful; they willingly comply with the experimenter's requests so that they can give the responses he is looking
for; and in attempting to cooperate with the experimenter, they actively search for cues which will give them some idea as to how they are supposed to behave in the experimental situation. With this kind of motivation on the part of the subjects, it is quite possible what the behavior experimenters observe is as much or more a function of this active problem solving by subjects as it is a result of the experimenter's conceptualization of the independent variables.

To illustrate the cooperative and complacent nature of the typical human subject, Orne describes his attempts to find a task so boring and ridiculous that subjects would refuse to do it. In one study he asked them to do a long series of addition problems which would obviously be impossible to complete within the day, and then to tear each sheet up as soon as it was finished. Subjects worked diligently at this job for several hours, until the experimenter finally gave up! Post-experimental interviews with the subjects revealed that they had (correctly) inferred that the experimenter must have had some legitimate reason for assigning such a task (i.e., many thought it was some sort of endurance test). In other studies (Orne & Evans, 1965), Orne describes subjects who would apparently do dangerous things like pick up a poisonous snake or a penny from fuming nitric acid, simply because the experimenter had asked them to do it.¹ In these ways, Orne's formulation seems quite viable.

¹The analogous observation of Milgram in the controversial obedience research (Milgram, 1965) is also demonstrative. Here subjects were told to administer shocks to a "victim" so severe that they thought they might have killed
Empirical Evidence for the Good-Subject Role

In general, there is widespread support for the notion that something like the process which Orne describes is part of the experimental situation, in that evidence for the good-subject role comes from a number of fields of research employing a variety of paradigms. The most clear demonstration of cooperative motivation is found in studies in which prior explicit information about the experimenter's expectations (usually by a confederate acting as another subject who has just finished the experiment) increases compliance with these expectations. For example, Levy (1967) had confederates inform subjects of the correct purpose of a Taffel-conditioning task ("... you have to sit there and make up sentences using these words she has on a card and it seems she wants you to make up sentences using I or We as the pronoun ... [p. 369]"). Subjects informed of this produced about one and one-half times more of the correct responses than did uninformed subjects. Levy's results have been replicated recently (Goldstein, Rosnow, Goodstadt & Suls, 1973; Smith, Helm, & Tedeschi, 1972) providing further support for the good subject in this paradigm.

A confederate was also the source of prior information in a study of the autokinetic effect (Alexander, Zucker & Brody, 1970) where the naive subject "accidentally" overheard another (the confederate) being run through the
experiment. This source of information about how much convergence was expected, together with other information directly delivered by the experimenter, was found to have considerable effect on the judgments of the real subjects. This provides additional evidence that subjects try to do what is expected of them, particularly when other cues to the correct response are ambiguous, as in an autokinetic procedure.

In a variety of other studies, strong support for the good subject has been demonstrated by the manipulation of cues which apparently increase the saliency of the experimenter's expectation, followed by a corresponding increase in cooperation with the hypothesis. For example, in a sensory-deprivation study, Orne and Schiebe (1964) found that the presence of a "panic button" produced hampered motor performance, as well as more reports of unusual experiences on the part of subjects than subjects not exposed to this cue. Similarly, Page and Scheidt (1971) varied the degree to which the presence of a weapon was pointed out, and found that subjects appeared much more aggressive under the high-salience condition (in line with their perception of the experimenter's expectations regarding the cue) than did subjects in the low-salience condition. Several other demonstrations like these have been reported with attitude-change studies. For example, Silverman (1968) found more attitude change from a classroom of subjects when he indicated that they were involved in a psychological experiment than when he intro-
duced the experimental task as part of a survey. In another study, Silverman (Silverman & Regula, 1968) found that subjects changed their attitude more in line with a message when a distraction (static on the tape subjects were listening to) was described as accidental. Apparently, subjects in the first condition assumed that the experiment was a test of their powers to concentrate and showed better attention than did the later subjects, again demonstrating that subjects try to do their best to live up to the experimenter's expectations. In other attitude-change studies, Rosnow (Rosnow & Suls, 1970; Rosnow, Holpner & Gitter, 1973) has shown that the use of a pre-test, in blatant form, enhances the good-subject role by communicating to subjects the notion of what the experimenter expects. Likewise the salience of the experimenter's expectations was manipulated by Rosnow, Rosenthal, McConochie, and Arms (1969) by presenting one-sided versus two-sided communications. The increased attitude change in the one-sided condition was interpreted as being due to the experimenter's expectations being much more transparent there. Finally, two conformity studies show that at times subjects might even risk looking irrational to please the experimenter. For instance, Bruehl and Solar (1970) varied the behavior of a confederate and found that conformity was greatest among those subjects who knew that the experimenter expected conformity. Similar results were reported by Adair (1972) in the description of a study by Allen (1966)
who found that conformity was greatest among subjects who were suspicious about the experimenter's hypothesis.

Of course, the strongest evidence comes from studies where cues are actually manipulated to increase awareness so the effect of this awareness on subject cooperation can be studied. But a number of suggestive findings from correlational studies also converge to support the notion that subjects try to help the experimenter find what they think he is looking for. In these studies, awareness is measured after the experiment, and the correlation between awareness and behavior is an indication of cooperation.

For example, in a figure-ground perception task, Page (1968) showed that subjects in a Schafer and Murphy perception task (Schafer & Murphy, 1943) perceived both aspects of an ambiguous figure, but subjects picked the portion which they thought the experimenter expected them to pick. Similarly, on a pursuit-motor task, subjects who thought the trials should improve performance did better than those who thought the trials should increase fatigue (Adair, 1970, reported in Adair, 1973). Page (1970) has also demonstrated that the typical communicator-credibility experiment is probably mediated by the good-subject role. When subjects read a message by a prestigious author, they were much more aware that the experimenter expected change, and they changed accordingly. Those subjects unaware of this hypothesis did not change. Subjects in a low-credibility condition (non-prestigious source) changed much less and were much less aware that the experimenter expected change (in line with
the actual expectations of the experimenter!).

Within this correlational methodology, a lengthy and complex controversy has enveloped many researchers studying verbal conditioning concerning the question of whether awareness of the correct reinforcement contingency is necessary for learning in the verbal-conditioning tasks of Taffel (1955) and Greenspoon (1955). Much of this controversy is centered on the many findings in which conditioning is highly correlated with post-experimental assessments of awareness of the correct contingency and experimenter expectations (Farber, 1963; Dulaney, 1962, 1968; Spielberger, 1962). Many agree now that it is unlikely that this controversy will ever be settled by empirical means (Greenspoon & Brownstein, 1968), probably because a crucial test is unlikely (Page, 1973); but the many studies in which awareness and conditioning are highly correlated (DeNike & Spielberger, 1963; Kennedy & Cormier, 1971; Monday, 1968; Page, 1967; Page & Lumia, 1968; Spielberger, DeNike, & Stein, 1965; DeNike, 1964; Dulaney, 1961, etc.) provide additional suggestive evidence for the good-subject role. In other studies with the paradigm, Holmes (1967) and Page (1970) have provided further evidence by showing that both awareness and conditioning are greater among subjects who have had more experience in psychology courses. This suggests that sophistication in psychology leads to awareness, which in turn produces more conditioning because the subject is motivated to help the experimenter and do what is expected of him. The same increase in
awareness and conditioning has also been found among volunteer as opposed to nonvolunteer subjects (Goldstein, et al., 1972).

In addition to the experimental and correlational investigations, a few other studies suggest more indirectly that Orne was correct, even though no manipulations or measures of awareness were performed. For example, Rosnow (1968) has suggested that demand characteristics may be responsible for the one-sided versus two-sided message effect. In a two-sided condition, subjects moved toward an anti-fraternity stand presumably because they thought this would be the position held by the faculty communicator. In the one-sided message condition, the expectations were apparently more clear, and subjects moved in the direction of the message. In a study of psychophysics (Juhasz & Sarbin, 1966) in which subjects were given distilled water and asked to give judgments of whether they tasted salt in the solutions, 83 percent of the subjects have at least one salt response, even though all solutions were distilled water. In a similar vein, Jacoby, Olson and Haddock (1971) found that subjects rated brand name beer samples according to expectations, even though the samples were all of the same beer. It is this sort of mounting evidence, both correlational and experimental, from a wide variety of paradigms, that have led many to doubt the interpretability of much of psychological research, from some studies concerned with cognitive processes (Neisser, 1967) to the treatment of insomnia (Eiseman,
1970), all on the basis of the good-subject role.

Other Roles

The evidence would seem to suggest that Orne was indeed correct and that subjects are unquestionably compliant in their response to cues which suggest an experimental hypothesis. However, not all researchers find that their subjects are so concerned with helping the experimenter. In fact, it has been suggested that there are some aspects about Orne, his laboratory, and his research that would give rise to an unusually large number of good subjects: first, he has been largely concerned with studying hypnosis, which is itself a somewhat cooperative process; secondly, he primarily used volunteer subjects who travel considerable distances in order to be in his research (Adair, 1973). There is a large amount of empirical evidence that volunteers are more perceptive and cooperative in an experimental situation than are non-volunteers (Goldstein, et al., 1972; Hood & Back, 1971; Horowitz & Gumenik, 1970; Rosenthal, 1965; Rosenthal & Rosnow, 1969; Rosnow & Rosenthal, 1970; Rosnow & Rosenthal, 1966; Rosnow & Suls, 1970; Rosnow & Aiken, 1973). Third, Orne has been described by his colleagues as an excellent researcher who takes great care in preparing his subjects (Adair, 1973) as well as possessing a dynamic personality and distinguished style which would enhance any cooperative inclination on the part of a subject (Shor, personal communication, 1972).
In contrast to Orne's formulation, others have described subjects who appear quite different than the "good subject." In fact, some have suggested that subjects will sometimes go out of their way to behave opposite to what they think the experimenter is hypothesizing. The comments regarding this "negativistic" role are not as clearly integrated as those regarding the good-subject role, and empirical evidence for it is even more scanty, but for various reasons these subjects have been designated as "recalcitrant" (Fillenbaum & Frey, 1970), and "negativistic" (Cook, et al., 1970), and the effects have been (rather indelicately) labeled the "screw you effect" (Masling, 1966) as well as the "boomerang effect" (Silverman, 1965). Speculation as to what causes negativism has ranged from the fact that subjects generally dislike situations in which they are treated as low-level employees (Argyris, 1968), that some subjects generally dislike experimental psychologists who attempt to manipulate them (Goldberg, 1965); that subjects dislike being forced to do experiments (Cox & Sipprelle, 1971; Black, Schumpert, & Welch, 1972); that they actively try to avoid situations with lessened freedom (Brehm, 1966; Rasnick & Schwartz, 1973), and that certain experimental treatments such as frustration can produce negativism (Silverman & Kleinman, 1967). It also appears that, if certain role expectations on the part of experimenters are violated, subjects become more uncooperative, such as in a recent study by Epstein, Suedfield, and Silverstein (1973) which showed
increased subject negativism when the experimenter was late for unexplained reasons.

While the proportion of negativistic subjects appears to be smaller, empirical evidence is growing that at least a small subsample of subjects do not necessarily follow the experimenter's expectations when they become aware of them. Aside from the experimental treatments of frustration, restricted freedom, and deception (Silverman, Schulman & Weisenthal, 1970; Stricker, Messick & Jackson, 1967) which have been shown to decrease subject cooperation with the experimenter's expectations, a number of studies have shown that even without these sorts of treatments, not all subjects who are aware of the experimenter's expectation do cooperate. For instance, in the many studies which demonstrate the high correlation between subject awareness and conditioning in both verbal and classical-conditioning paradigms (DeNike, 1964; DeNike & Spielberger, 1963; Dulaney, 1961; Holmes, 1967; Kennedy & Cormier, 1971; Monday, 1968; Page, 1968, 1969, 1970; Page & Lumia, 1968; Spielberger, 1962; Spielberger, DeNike & Stein, 1965), not all aware subjects do condition.

In general, the literature suggests that the incidence of the cooperation far outnumbers the negativism, as, for example, the study by Kennedy and Cormier (1971) which showed that only 5 out of their 108 subjects were highly aware of the contingencies and hypothesis, but had negative behavioral intentions and negative pre-experimental attitudes toward experimental psychology. While the number is not
great, it appeared to these researchers to provide incipient fulfillment of Rosenthal's (1966) prophesy that a trend may be materializing where subjects, more knowledgable about classic research than preceding generations, are determined to show experimenters that they are not one of those "mindless acquiescers which instructors of elementary psychology courses are likely to teach about" [p. 11, 1966].

The importance of subject motivation to cooperate in the verbal-conditioning paradigm, and the wide opportunity for it to be decreased by interpersonal variables, is also shown in a lengthy review by Kessel and Barber (1968) of the various interpersonal treatments which lowers subject motivations, such as unrewarding prior interaction and perceived dissimilarity between the experimenter and subject (Sapolsky, 1960).

Other research suggests that mere information about the experiment decreases rather than increases conformity (Adair, 1972; Glinski, Glinski, & Slatten, 1970; Horowitz & Rothchild, 1970). While a satisfactory explanation of the contradictory findings remains to be adequately resolved, it has been suggested (Adair, 1972) that the effect which information will have depends upon the subject's perception of the conformity as being irrational dependency or good-natured cooperation. That the effect of information in the conformity paradigm is complex is also demonstrated by another study by Gallo, Smith and Mumford (1973) where either complete or partial information produced no differences from a group with no information. What it is about the different procedures and information in these studies which produces these discrepant effects across laboratories remains an important
problem to be solved, but it would seem that Adair's suggestion would provide a good lead. This notion underscores the subject's perception as an important link between the independent variable and the dependent variable in the laboratory study, as well as emphasizes the importance of evaluation apprehension, a topic to which we now turn.

**Evaluation Apprehension and the Defensive Subject**

The suggestion that the effects of cues in the situation on subject cooperation depends upon the evaluation of cooperative behavior as to its socially-desirable or undesirable qualities brings us to the third often cited subject role. This role is based on Rosenberg's (1965) notion of evaluation apprehension, which is defined as "an anxiety-toned concern [on the part of a subject] that he win a positive evaluation from the experimenter, or at least that he provide no grounds for a negative one [p. 29]."

Rosenberg's explicit formulation is recent, but the notion that people are very much concerned with putting their best foot forward or presenting themselves in a socially-desirable manner during a testing situation has been with us for some time (Edwards, 1957; Reicken, 1962; Rosensweig, 1933). Indeed, many of our deception procedures such as cover stories, filler items, and lie scales manifest psychologists' beliefs that, if given the opportunity, the subject will conceal or exhibit, exaggerate or belittle those qualities he believes will be positively or negatively
appraised in an experiment. This role, that of the apprehensive subject, has been considered as probably the most important one by a number of researchers (Berkowitz, 1971; Silverman, 1968; Weber & Cook, 1972) and so a description of its original inception is given at this point.

Evaluation apprehension was originally proposed by Rosenberg (1965) as an alternative explanation to a classic experiment in cognitive dissonance (Cohen, 1962) in which subjects were offered various monetary rewards to write a counter-attitudinal essay. In line with the prediction from dissonance theory that the most attitude change would occur under conditions with least external justification, the students offered a small amount were more positive in their appraisal than were those given a large amount.

To Rosenberg, however, it seemed more plausible that, rather than creating differential levels of cognitive dissonance, what the independent variable of monetary amount did was to create different levels of evaluation apprehension. According to this view, the large reward offered to the subject in the low-dissonance condition aroused his suspicion that his honesty in resisting bribery was being tested. The subject, in Rosenberg's words, would be likely to reason that the experimenters

"... probably want to see whether getting paid so much will affect my own attitude, whether it will influence me, whether I am the kind of person whose views can be changed by buying him off." The subject who has formulated such a subjective hypothesis about the real purpose of the experimental situation will be prone to resist giving evidence of attitude change: for to do so would,
as he perceives it, convey something unattractive about himself, would lead to his being negatively evaluated by the experimenter [p. 286, 1969].

When Rosenberg replicated the Cohen study, but separated the dissonance and measurement procedures by making the attitude survey and essay writing appear to be parts of two different experiments, the opposite effect was found. That is, subjects who were paid the most money showed the most agreement with the counter-attitudinal statements. Thus it appears that the original data were produced by an artifact of elicited evaluation apprehension, and when such apprehension was minimized, the "dissonance effect" was not found.

The Rosenberg re-explanation of the dissonance study has received the attention and controversy one might expect (e.g., Cook, 1969; Jones, Cooper, Carlsmith, Collins, & Helmreich, 1966; Linder, Cooper, & Jones, 1967) and failures of others to replicate the Rosenberg findings have produced even more complexities (see, for example, Jones & Gerard, 1967, pp. 494-496 on when commitment occurs). However, for the purposes of the present discussion, what is important is the formulation of evaluation apprehension as a primary concern of human subjects in the laboratory and its resulting effects on behavior.

A recent study by Alexander and Knight (1971) seems to provide some additional empirical support for evaluation apprehension as a mediator between the experimental situation of the dissonance study and the eventuating behavior as proposed by Rosenberg. The Alexander and Knight study was
essentially an interpersonal simulation (Bem, 1967) of the experiment, in that Alexander and Knight's subjects listened to a tape recorded description of a typical subject who was described through the procedures of a dissonance study by Festinger and Carlsmith (1959) in which subjects were given various amounts for lying about an experiment to another subject. When the Alexander and Knight subjects were asked to make character inferences about a typical subject who did change his attitude in the high-reward condition, they supported Rosenberg's argument by seeing him as less honest and more materialistic than a subject who changed his attitude in the low-reward condition. Thus Alexander and Knight concluded that it was very plausible that the subjects in the original study were behaving such that their "concern to maximize desirable identity impressions determined the dependent variable . . . [pp. 74-75]."

Other studies have shown the facility with which subjects make evaluations based on experimental behavior. Kauffman (1971) has essentially replicated Alexander and Knight's procedures and findings in another counter-attitudinal study and Alexander and Weil (1969) have demonstrated the notion that subjects act in line with the most favorable inferences in a Prisoner's Dilemma Game.

While these studies, particularly the Alexander and Knight study, provide evidence that Rosenberg's notions of evaluation apprehension do operate in these paradigms, more direct data come from Rosenberg himself (1969), however, in
his programmatic approach to the study of the antecedents and effects of this motive. Representative of his studies was a person-perception task in which slides of faces were viewed by subjects who were asked to rate them according to how much they liked them and how successful they think the people owning the faces were. This task was given with a variety of instructions, among which was usually a "Background Information Sheet" which delivered various evaluation cues. One of these was a general cue, designed to heighten any evaluation apprehension that the subject may have had as he entered the experiment. Thus the subject was told in the past, performance in this task had been shown to correlate with personality, particularly with whether the rater was psychologically "mature" or "immature." In conditions designed to lower general evaluation apprehension, the subject was at this point informed that the purpose of the study was to merely construct normative data against which later comparisons would be made. Also given in the Background Information Sheet was information which provided directional evaluation cues, that is, cues which told the subject which behavior would elicit the most favorable evaluation from the experimenter. Thus, the group given a directional cue was also told that the main burden of past research (with various invented journal articles cited) was to show that people who are more psychologically mature and healthy show a greater liking for strangers than do immature people.
Another group was told the opposite, that mature people generally show a greater disliking.

The pictures are then rated on a 21-point like-dislike scale, and the results showed that the picture ratings were reasonably consistently affected by the evaluation cues. It is also notable that subjects who scored high on the Marlowe-Crowne Social Desirability Scale (Crowne & Marlow, 1960), a scale which measured need for social approval, were more influenced by evaluation cues than were those who scored lower.

Rosenberg has extended this research to other procedures, including a perceptual-motor task, and to the study of independent variables which interact with evaluation-apprehension cues, such as partial versus complete feedback, and whether or not the experimenter holds special prizes for the subject (Rosenberg, 1969). For the most part, the results of his research program were consistent in demonstrating that 1) such apprehension can be exacerbated by situational cues, 2) subjects bias their behavior in line with favorable evaluations suggested by these cues, and 3) subjects high in need for social approval are affected more by these cues.

Other support for the notion that subjects try to get favorable evaluations from the experimenter come from a variety of sources. One particularly clever demonstration was by Gustafson and Orne (1965) where subjects were put in a lie-detector apparatus and told either that only psychopaths can successfully deceive in such a situation, or that mature
and intelligent persons are able to deceive. Subjects were then told that they were successfully deceiving if they were in the psychopath condition, or unsuccessful at deceiving if they were in the intelligence condition. Skin resistance then increased as subjects learned about their failure to produce socially-desirable responses.

Other evidence from experimental manipulations appear in the verbal-conditioning paradigm where Page (1971) has demonstrated that the administration of a test, which is taken from the MMPI and includes general adjustment items and some items dealing with conformity behavior, significantly reduced the amount of correct responses subjects gave. Apparently, according to Page, the personality test provided the kind of evaluation cues similar to Rosenberg's Background Information Sheet which suggested that the experimental task measured personality (the MMPI general items) and that conditioning was a measure of conformity (suggested by the conformity items). Page interpreted this finding as demonstrating that cooperation in such an experiment depends upon low levels of evaluation apprehension. This is a point about which more will be said later, but it should be noted that correlational evidence from Katkin, Risk, and Spielberger (1966) also supports the contention that subjects will not show conditioning if they think it demonstrates conformity. In this study subjects who showed negative behavioral intentions but full awareness, often attributed conditioning to conformity and reported that they purposely tried not to
conform. This occurred especially when the experimenter was seen as a high-status professor, rather than a low-status undergraduate.

Several studies in the attitude-change literature further elaborate the role of evaluation apprehension through direct manipulation. One of them is by Silverman (1968) who found that the greater attitude change found in the context of an experiment as opposed to a survey occurred principally in the condition in which subjects were requested to sign their names to their responses. Apparently, the name signing decreased anonymity, thereby increasing evaluation apprehension and hence, susceptibility to the demand characteristics.

Another suggestive study was done by Silverman (Silverman, Schulman, & Wiesenthal, 1970) who varied the degree to which subjects were deceived in prior experiments and then gave them a series of standard psychological measures. The deceived subjects showed compliance with perceived demand characteristics in some cases, such as in a persuasion test, but less compliance with demand characteristics in a semantic-differential rating task. In another test, deceived subjects gave more favorable self-descriptions. These results were synthesized to suggest that the deception enhanced evaluation apprehension, leading to more careful self-presentation, less cooperation, but more persuasion, since the messages were abstracts of published articles written in a logical and factual style, and agreement with
them would presumably show open-mindedness. Unfortunately, the study does not include any evidence that the subjects actually perceived the tasks as such, and so the results are only suggestive.

While these studies have been explicitly interpreted in terms of Rosenberg's formulations, it seems that a wide variety of other phenomena could also be explained with respect to evaluation apprehension. For instance, recent evidence shows that it seems to be at the heart of the social facilitation effect, whereby the presence of others enhances dominant responses while inhibiting subdominant ones (Zajonc, 1965, 1966). The essential conditions for this effect have received much theoretical and empirical attention (i.e., Cohen & Davis, 1973; Cotrell, 1968; Cox, 1968; Hartens, 1969; Matlin & Zajonc, 1968; Zajonc, 1969), but an important aspect seems to be the evaluative nature of the situation (Cotrell, 1968; Jones & Gerard, 1967). For example, Henchy and Glass (1968) have shown that the social facilitation effect depends upon an observer being perceived as an expert who makes careful observations of the subjects' performance, even if not physically present. When the observer was perceived as a student just watching a psychology experiment, the facilitation did not occur. Similar findings are reported by Cotrell, Wack, Sekerak and Rittle (1968) and Paul and Murdock (1971) whose subjects showed the facilitation effect with or without an audience only if they anticipated later evaluation.
In an even broader vein, Weber and Cook (1972) have suggested that with few exceptions, most of the literature on subject roles can be interpreted in terms of the operation of evaluation apprehension. For example, while subjects' behavior in verbal-conditioning studies suggest the good-subject role, cooperation in this situation would also show the experimenter they had discerned the reinforcement contingency, and solved the problem of what was making the experimenter say "good." This would demonstrate some degree of intelligence. Likewise, in studies of attitude change (as Silverman [1970] has noted) cooperation with demand characteristics can be interpreted as demonstrating open-mindedness and flexibility. To the extent that subjects feel that compliance with expectations would reflect negatively upon them, they react negativistically, such as in the conformity studies of Horowitz and Rothchild (1970) and Adair (1972). These observations have led Weber and Cook to conclude that evidence for the good and negativistic roles are consistently confounded with evaluation apprehension, and that the most parsimonious formulation of subject motivation may be that evaluation apprehension is responsible for all subject-role behavior.

**Evaluation Apprehension versus the Good Subject:**

**Situational Cues**

At this point, it seems that much of the current literature on subject motivation is converging to support
Weber and Cook's conclusion that the good subject may be a special instance of the more general and pervasive evaluation apprehension motive. This position contrasts sharply with Orne's statement on the question:

Admittedly, subjects are concerned about their performance in terms of reinforcing their self-image; nonetheless, they seem even more concerned with the utility of their performance [1969, p. 778].

The general term, demand characteristics, is open to considerable classification and specification, and it seems likely that the relative importance of cooperation versus apprehension might very well depend upon the different kinds of demand characteristics in any given experimental situation. For example, it seems plausible that the good subject is dependent upon the perception of hypothesis cues, those which would help him to discern what it is that the experimenter is expecting. These hypothesis cues include stimuli which Orne originally specified as demand characteristics in his primary emphasis on the good subject as the chief subject motivation. From the review of the good subject literature, it seems that hypothesis cues operate in a wide variety of forms, from "panic buttons" to brand names of beer in taste-sensitivity experiments, to pre-tests in attitude-change studies. These hypothesis cues are relatively obvious in nature, but Rosenthal's research also suggests that they can be communicated along very subtle lines from the experimenter, including the number of glances he gives while reading instructions, or the relative vocal emphasis he gives to certain words.

The second major category of demand characteristics are then evaluation cues which serve to elicit or direct evaluation apprehension, by suggesting to the subject that his personality or intelligence is under scrutiny. In order for the subject to behave apprehensively, he has to perceive these cues, which can be further subdivided into general cues and directional cues. General cues elicit generalized evaluation apprehension without the specific information of how to monitor one's performance to receive the best evaluation from the experimenter. They have been shown to be operative in such forms as a camera lens (Henchy & Glass, 1968) and the presence of an experimenter (Schulman, 1967) and would also seem to be communicated by an experiment title which emphasizes the study of personality (Page, 1973; Silverman, 1973) or by awesome machinery in the experimental laboratory (Franks & Jenkins, 1968).

---

^2 Orne (1969) points out, however, that while experimenter bias effects can be included under the notion of demand characteristics in the sense that they are mediated to the subject by demands (Adair has also drawn this relationship, 1973), in another sense they are quite different. The experimenter-bias effect is rooted in the motives of the experimenter, whereas demand characteristics are conceptualized from the subject's point of view. Experimenter-expectancy effect can invade all science, as Rosenthal has demonstrated with "N Rays" (Rosenthal, 1966), but demand characteristics are only relevant to research with human subjects, as light rays and other physical entities do not guess the hypothesis and perform accordingly. However, given the central issues of the present research, that of the human subject and his motives, it seems feasible for the present purposes to include the experimenter-bias effects as a form of hypothesis cues, under the general rubric of demand characteristics.
Empirical Evidence

With this scheme in mind, it is possible to look at the relevant literature in terms of the interaction of hypothesis cues and either kinds of evaluation cues on behavior.

For example, research dealing with the general cues has appeared within the domain of the experimenter-expectancy effect (hereafter EEE), following Rosenthal's (1966) suggestion that some minimal amount of evaluation apprehension may be necessary for the EEE to operate at all. This notion has recently been supported empirically by Minor (1970), who introduced general evaluation cues by informing one group of subjects that the purpose of the person-perception task they were about to begin was to replicate the previous finding that those who are very inaccurate in their perceptions of people are maladjusted. This served as the High Evaluation Apprehension condition. The Low Evaluation Apprehension condition consisted of telling another group of subjects that the purpose of the perception task was to gather baseline data against which the effects of fatigue and practice would be compared. Thus generalized evaluation apprehension was lowered by the information that personal characteristics were not relevant to the task. When experimenters led to expect different ratings then tested these subjects, it was found that the EEE was only operative in the general cue condition. Apparently, when evaluation apprehension was lowered, subjects either did not perceive the experimenter's cues, or did not
bother to comply with them. This finding was also reported by Johnson (1973) who used a marble-dropping task and suggested to some subjects that the rate of dropping was related to intelligence. The EEE was only found with these subjects, and only with experimenters who showed some concern about the outcome, suggesting that the EEE is not only mediated by evaluation apprehension on the part of the subject, but possibly also by apprehension on the part of the experimenter as well.

To the extent that the psychological experiment is itself a general-evaluation cue, one would expect, with Rosenberg, that a certain amount would operate without any special manipulation of evaluation cues. That this certain amount is enough to mediate the EEE has been recently demonstrated by Duncan, Rosenberg, and Finkelstein (1969) who extended Minor's study to include a control group which was not told anything regarding the nature of the task. The results showed that subjects were influenced by the experimenter cues in both the High Evaluation condition as well as when no cue was given, but when it was suppressed in the Low Evaluation condition, the EEE again failed to be demonstrated.

What these studies suggest is that, unless evaluation apprehension is explicitly decreased, it will operate at a high-enough level to mediate subtle demand characteristics to produce the good subject. But because the operation of the experimenter cues depends upon the level of evaluation apprehension, it would appear that subjects are only "good" if
they believe they are to be evaluated. This position has recently been expressed with respect to less-subtle hypothesis cues, such as in attitude-change studies (Silverman, 1965, 1970) and the effects of aggressive cues on aggressive behavior (Berkowitz, 1971). It seems that the current Zeitgeist is viewing the apprehensive subject as much more common and basic than the good one.

The second line of evidence which has led to this position comes from those who have been engaged in directly testing the relative importance of the good versus the apprehensive roles by pitting directional evaluation cues against hypothesis cues. This was first done by Sigall, Aronson, and Van Hoose (1970) who attempted to present their subjects with a dilemma: they could either fulfill the experimenter's expectations, or they could maximize looking good, but they could not do both. Since this study has been described as "crucial," (Rosnow, 1970; Weber & Cook, 1972) in demonstrating the prepotency of evaluation apprehension over cooperation, a close look at it is important.

In their effort to pit the good and apprehensive motives against each other, these investigators gave subjects an assignment to copy telephone numbers for seven minutes, followed by further instructions regarding the task, and then a second trial on the task. In all conditions, the second trial was accompanied by a change in the lighting of the room, in that in all cases, illumination was cut in half. The experimenter mentioned to all subjects that he was investi-
gating a hypothesis concerning the effect of illumination on work performance.

The information given between trials comprised the experimental treatment. In the "Increased-Output" condition, subjects were told by the experimenter that she expected an increase of twenty numbers for the next trial. Thus subjects were given the experimenter's hypothesis in explicit form, which was to increase. In the "Decreased-Output" condition, subjects were told that the experimenter expected a decrease of twenty numbers during the second trial. In both the "Increased Output" condition and the "Decreased Output" condition, it was assumed by the experimenter that subjects would implicitly understand that copying more numbers was a better performance, and so any evaluation apprehension would lead to increased performance. Hence, in both conditions, there was an assumed directional evaluation cue to increase. This was not the case in their "Decreased-Output-Obsessive-Compulsive" condition in which subjects were told that "people who feel compelled to rush at a trivial, boring task like copying numbers tend to be obsessive compulsive [p. 274]." These Obsessive-Compulsive subjects were also informed of the hypothesis to decrease. (Unfortunately, there was no Obsessive-Compulsive group with a hypothesis cue to increase.)

The results of this study demonstrated that information provided after the first trial systematically changed behavior on the second trial. Whereas there was no significant change in performance for a control group which merely
did two trials without any intervening information, the other groups did vary their performance according to the demand characteristics communicated. In both conditions where the evaluation cues and hypothesis cues were congruent, subjects changed their performance appropriately. That is, subjects in the Increased-Output condition increased, and those in the Decreased-Output-Obsessive-Compulsive condition lowered their performance. In the critical condition where the cues were conflicting, subjects increased their performance, even though the experimenter told them she expected a decrease. Because of the performance in this condition, the authors concluded that subjects, when faced with the dilemma of either confirming the hypothesis or presenting themselves favorably, will choose the latter. Therefore, the apprehensive role is more important than the good-subject role. Thus the results converge with those mentioned above regarding the EEE and suggest that subjects will only cooperate with the experimenter when it will enhance their self-presentation.

The conclusion seems plausible in terms of the data reported, and because it fits so well with these other lines of research, it was somewhat unexpected that a further elaboration of this research paradigm suggested that the conclusions may have been quite unwarranted. The essential weakness of the study (which has been characteristic of most of the research reviewed above) was a failure to adequately perform manipulation checks to see if the cues in the situation really did lead the subject to perceive the experiment in the
manner which the investigators assumed they would. Sigall et al. did perform some very indirect attempts to check their manipulations, such as another control group which was used to verify the evaluation cue. This group was not given any hypothesis, but told only that the experimenter was interested in the effects of illumination on performance. On the second trial, this group increased its performance, suggesting to the authors that the directional evaluation cue was operating because subjects were trying to impress the experimenter with better performance. However, this was still quite inferential, not being based on subject reports. They also used a post-experimental questionnaire to weed subjects who might have been aware of the real hypothesis (i.e., the effects of cues) but in no case was there an attempt to validate the adequacy of the hypothesis cue in generating differential perceptions. Without knowing whether the cues were really creating differential awareness, there is little that can be said about the nature of subject motives to respond on the basis of them.

Fortunately, Adair and Schachter (1972) dealt with this problem by replicating and extending the Sigall et al. study. The primary rationale behind the Adair and Schachter experiment was their argument that even though the performance of Sigall et al.'s Decreased and Increased Output groups appeared the same (i.e., both groups increased their performance), there is no evidence that the subjects behaved on the basis of the same perceptions, and thus the same evaluation
motivation. Was performance in the critical Decreased Output condition a result of the over-riding evaluation apprehension motive as Sigall et al. concluded, or could it have been due to a failure to communicate the hypothesis cues adequately? To test this question, Adair and Schachter varied the explicitness with which the hypothesis cue was communicated. In their Implicit condition, they used the exact procedure of Sigall et al. by informing subjects that the experimenter expected a decrease on the second trial. In the Explicit hypothesis conditions, this expectation was then elaborated by explaining the reasons for the hypothesis. Thus each subject was told that "we have a theory that due to fatigue arising out of performing the practice trial, your output on this task would be expected to decrease [p. 78]."

The results of this explicit-implicit manipulation were striking. Whereas those subjects in the implicit condition again increased their performance, those in the explicit condition decreased. The explanation for these results becomes clearer with the consideration of essential data from the post-experimental questionnaire. Here subjects were asked "How do you think you were supposed to perform on the second trial?" and they were given an opportunity to check one of the three responses: Increase (scored as 1), Stay the Same (scored as 2), or Decrease (scored as 3). These scores, when subjected to an analysis of variance, indicated that there was a significant effect of the communication variable. Whereas subjects in the explicit condition had a
mean score of 2.38, those in the Implicit (Sigall) condition had a mean score of 1.18. These means indicate that, although the subjects in the explicit condition were not impressively uniform in their perception of the hypothesis to decrease (uniform perception would have yielded a mean score of 3.00), they were at least more accurate than the subjects in the Implicit condition who showed a greater tendency to perceive a hypothesis to increase. Thus, it seems likely that the Sigall et al. subjects in the Decreased-Output condition were not sacrificing the experimenter's expectations to create a good impression. It is possible that they were behaving in line with what they thought he was hypothesizing. Thus, we are left with no clear-cut support for the apprehensive role over the good-subject role.

Unfortunately, there still is not clear-cut support for the good-subject over the apprehensive one, either, for the Adair and Schachter procedure had a serious problem which prevented the possibility of it being a reasonable test. True, subjects in the explicit conditions decreased their performance in line with experimental expectations. But by making the hypothesis explicit in the manner they did, it would seem likely that they destroyed any evaluation cues.

The casualness with which the evaluation cue was considered was started by Sigall et al. when they assumed that the more numbers copied, the more favorable the evaluation the subject would perceive he could win from the experimenter. What evidence there was for this conclusion was indirect and
tentative; it was based primarily on a group which was not
given a hypothesis cue, but when told that the experimenter
was interested in the effects of illumination on performance,
increased their performance. The Adair and Schachter study
seemed to buy this assumption by leaving the evaluation cue
as it was—basically more implicit than the original hypothe-
sis cue. However, in Adair and Schachter's explicit condi-
tion, they told subjects that performance was expected to
decrease because of fatigue from the first trial. It is
reasonable to assume that given this explanation in terms of
fatigue, any possible perception of an evaluation cue would
be seriously jeopardized. Under these conditions, any
decrease in performance is expected because of situational
determinants, i.e., a difficult task which makes all people
in that condition tired. Either behavior is attributed to
the situation, or it is attributed to personal qualities, a
basic assumption underlying recent theoretical developments
in the attribution processes (Heider, 1958; Kelly, 1967).
Therefore, conforming to the experimenter's hypothesis would
not have cost the subject anything in self-presentation
because the experimenter had already given him a logical
explanation for his behavior which did not reflect negatively
upon him. In fact, this was precisely the information (i.e.,
"we are studying the effects of fatigue") which was used in
previous studies to create a Low Evaluation Apprehension Con-
dition (Duncan, Rosenthal, Finkelstein, 1969; Minor, 1970;
Johnson, 1973). Unfortunately, Adair and Schachter provide
us with no evidence that subjects perceived any evaluation cue, and so we are left with the very plausible interpretation that there may not have been any.

In order to adequately test this question, manipulation checks are essential. What is needed is an experimental manipulation of cues which converge with awareness data to adequately determine the relative importance of these motives.

For this reason two recent attempts at separating these roles are also questionable. Rosnow, Goodstadt, Suls, and Gitter (1973) assumed the communication of the experimenter's hypothesis in an impression-formation task and pitted it against information from a confederate who, in the middle of the experiment, suggested directional evaluation cues. Even though the authors concluded that the importance of the evaluation apprehensive subject was demonstrated over the good-subject role, this is not clear because 1) no assessment of the perception of the cues was provided, and 2) their dependent variable data was not entirely consistent with their predictions.

In the second study, Geller and Endler (1972) based their independent variables on post-experimental data for hypothesis awareness and suspicion in a conformity task, and concluded that the apprehensive subject reigned over the good one because only those in a High Hypothesis Aware but Low Suspicion group conformed. However, their evidence is questionable on several counts. First, without any manipulation of the situational cues, it is difficult to know just how subjects sorted themselves into the post-experimental
awareness categories. Without experimental manipulation, a serious subject selection factor may operate which prohibits clear inferences regarding the causative role of subject motives. Secondly, the separation of the suspicion from hypothesis awareness in the conformity paradigm is unclear, since they would seem to be highly related. Indeed, the cell frequencies were grossly uneven. This situation violates the assumptions of the ANOVA model, and throwing out two-thirds of the data in two different cells, as they did, resulted in a great loss of information. Third, as discussed previously, literature on the conformity paradigm is unclear regarding the perception that subjects have of the evaluation of conformity behavior. In some cases it appears that awareness increases conformity, whereas others show the opposite or no effect. Without a direct measure of what conformity means to the subject, it is impossible to logically infer anything about the effects of that evaluation of motivation. A better study would be within a paradigm in which the relationship between the subject's awareness of the hypothesis and his cooperation was reliable.

Thus, the empirical support for the proposition that evaluation apprehension is more important and basic than the good-subject role is inadequate. Furthermore, it seems likely that for several other reasons, the apprehension role has been given an undue amount of importance in the analysis of subject motivations.
First, as Weber and Cook (1972) point out, it is difficult to invalidate. The apprehensive role has very wide post-hoc explanatory latitude, as it can be evoked to explain a great deal of subject behavior in many different paradigms. However, it is not clear that in the many studies in which the good and apprehensive roles would lead to the same behavior, such as those involving problem solving, ability tasks and attitude change, that it is a more important determinant than is the subject's desire to cooperate.

Secondly, the studies of the experimenter-expectancy effect which highlight the apprehensive role by suggesting that some minimal amounts of evaluation apprehension is necessary for the effect to operate, represent a special situation in which it is likely that the importance of evaluation motives would be unduly inflated. This arises from the fact that in this paradigm, the hypothesis cues which the subject must follow if he is to demonstrate his good role are extremely subtle and difficult to recognize (indeed, experimenters are hardly able to). In this circumstance, it seems an unfair test of the good versus apprehensive roles because the subject must go beyond the obvious to discover the necessary hypothesis cues. It is entirely possible that in these situations, only those who are motivated by some other concern, such as evaluation apprehension, might be willing or able to do this. The research doesn't tell us much about the wide range of experiments in which the hypothesis cues are more obvious to the subject. In these cases, it would seem reasonable that, while the presence of general evaluation cues might enhance compliance with hypothesis cues, subjects would also cooperate
with them even when general evaluation apprehension was not purposely enhanced. If this were the case, the good-subject role would have status beyond a mere particular of the more general apprehensive role. What is needed is a paradigm where the hypothesis cues are obvious enough to be perceived without special cues for evaluation apprehension.

By selecting a paradigm in which the awareness of hypothesis cues is reliably correlated with cooperation without special manipulations of evaluation cues, the effects of evaluation cues can be studied independently of hypothesis awareness. This would enable a factorial design in which hypothesis cues can be crossed with evaluation cues in an orthogonal manner appropriate for the analysis of variance model so that the nature of any interaction can be assessed.

The classical conditioning of attitude experiments, originally developed by Staats and Staats (1957) was chosen because it seemed to have characteristics suitable for examining cues which elicit cooperation and cues which elicit evaluation apprehension.

The Classical Conditioning of Attitudes Paradigm

In this task, subjects are brought into the laboratory in small groups (5 to 10) and told that the experiment deals with intermodality learning. They are instructed that nonsense syllables will be flashed up on the screen in front of them, and that as this is done, the experimenter will read
words. Subjects are instructed to repeat the meaningful words as they are read. After this procedure, each subject is handed a rating booklet and asked to rate each nonsense syllable in the booklet according to how pleasant or unpleasant he feels it is (see Appendix D for the booklet). The rating scale is a 7-point semantic differential type dimension which runs from Pleasant to Unpleasant.

It has been demonstrated many times (Cohen, 1964; Insko & Oakes, 1966; Hare, 1964; Staats & Staats, 1957; Weber & Riddell, 1973) that those nonsense syllables which are flashed while the experimenter reads positive words (such as sunshine, pleasant, rich, and healthy) are afterwards rated toward the pleasant end of the dimension, and those syllables which are paired with negative words (such as ugly, enemy, dirty, worry) receive more negative ratings. This was originally interpreted as demonstrating the classical conditioning of attitudes (Staats & Staats, 1957) because the nonsense syllables (CS) appeared to take on the meaning (CR) of the meaningful words (UCS) when they were paired together. In line with the classical-conditioning explanation, Staats and Staats argued for the elicited nature of this behavior, because they found the effect to hold up even when they discarded the data of the few subjects who, at the end of the experiment, were able to answer the question "What did you think the experiment was about?" with some indication of having understood the conditioning process. Thus, it appeared to the Staats that the effects held with subjects who were
completely unaware of being conditioned.

However, as one of the more consistent findings in the psychological literature, there has been the repeated demonstration that in this task, "conditioning" is highly correlated with post-experimental reports of awareness, if the subject is given a reasonable opportunity to report it (Cohen, 1964; Hare, 1964; Insko & Oakes, 1966; Page, 1969; Weber & Riddell, 1973). This has been taken as evidence that the awareness of the contingencies (what nonsense syllables were paired with positive and negative words) and awareness of the hypothesis (that consequently the experimenter expects one to rate the nonsense syllables accordingly) mediate behavior in this experiment, rather than the unconscious mechanical effects of conditioning trials.

Although Staats has subsequently (1971) argued that the demonstration of this correlation doesn't rule out the conditioning interpretation, evidence suggesting the alternative demand awareness explanation is steadily growing. For example, Page has shown in a series of studies that the so-called conditioning effect is facilitated by subject's sophistication in psychology, and hampered by making the demands more difficult to discern by the addition of filler syllables (1969; 1970). Recently, he has also shown that subjects can easily produce the opposite response if requested (1973), further supporting the view that the responding in this situation is more likely a product of compliance with experimenter expectations than it is the elicited response of
classical conditioning.

Unlike the number-copying task which was used by Sigall, Aronson and Van Hoose for studying the effects of subject roles, the classical conditioning task has a rich history of controversy and empirical attention. It also has several other characteristics which make it particularly suitable for the purposes of the present study. First, unlike the Picture Rating task used to study the role of evaluation apprehension (Rosenberg, 1969) and its effects on mediating the good-subject role (Duncan, Rosenberg, & Finkelstein, 1969; Minor, 1970) the classical-conditioning task has, inherent in its procedures, hypothesis cues which are much more obvious. This is in part because it employs a kind of repeated measures design in which the subject sees positive, negative, and neutral words paired with different nonsense syllables, making it relatively easy for him to discern the pattern (about half of the subjects do, as shown by Page, 1969, and Weber and Riddell, 1973). Once the pattern is discerned (i.e., that YOF is always paired with positive words and WUH with negative) the hypothesis cue is almost always perceived. Thus, as Orne (1969) has suggested, studies in which the subject gets more than one level of the independent variable are especially prone to the influence of demand characteristics because they facilitate the perception of what the experimenter expects on the basis of what changes in the experiment. This seems to be particularly the case in this paradigm.
By enhancing the awareness of the contingency through instructing subjects of the relationship before the conditioning trials, hypothesis awareness can be enhanced and the results on conditioning and reports of awareness assessed. Independent of this manipulation, evaluation cues can be introduced which would suggest a negative evaluation from the experimenter. The effects of these evaluation cues can be assessed in their interaction with hypothesis cues on the dependent variable, and on the measures of awareness. This then would give the necessary situation for the most valid study of subject roles to date—the independent manipulations of hypothesis and evaluation awareness in concert with their effects on subject awareness and cooperation.

Thus, the purpose of the present research was to adequately examine the operation of the good and apprehensive roles by independent manipulation of hypothesis and evaluations cues. The effect of these manipulations could then be discerned on the dependent variable, conditioning, which, when accompanied with awareness, would indicate cooperation. The independent manipulation of these two types of demand characteristics would also contribute important refinements to some of the very vague constructs suggested to account for subject behavior. For example, the term "demand characteristics" can be further differentiated into two major types, hypothesis and evaluation cues. Also, the term "subject role" is one which has been used often (Adair, 1973; Weber & Cook, 1972) without clear definition. In the context of the present
research a "subject role" is operationalized as the effects of these types of cues on behavior. Thus, the good subject would receive support if the presence of hypothesis cues led to the increase in the hypothesized behavior. The apprehensive role would be empirically substantiated to the extent that evaluation cues enhanced the positively-valued behavior or, in this case, decreased the negatively-valued behavior.

It was predicted that the degree of subject cooperation would depend upon the specific number and type of these cues present in a situation, and that the question of which role is more basic becomes less important than the knowledge of what cues in the experiment interact to elicit each role.

In order to provide further convergence for the central hypothesis, an individual-difference measure was chosen for study that should theoretically relate to the process of perceiving and being affected by evaluation cues. Since the Social Desirability (SD) scale of Crowne and Marlow (1964) was designed to measure need for social approval, and has been used in past studies of evaluation apprehension (Rosenberg, 1969) it was administered in order to test the prediction that subjects high in need for social approval would be more affected by the evaluation cues than subjects low in this need. Empirical support for this hypothesis would further validate the evaluation cues and provide a more complete verification of the central hypothesis, through anchoring the process to a set of responses not collected in the experimental situation.
The Study of Post-Experimental Awareness

As suggested by the preceding discussion, subject reports of awareness are critical to the study of the effect of demand characteristics on behavior. Without the validation of such cues in the phenomenological perspective of the subject, it is inappropriate to draw conclusions about the effects of various cues on subject cooperation.

Subject awareness has also become a central issue in a continuing controversy within the conditioning literature, where a longstanding question concerns whether subjects must be aware in order to condition. This controversy represents a basic dissenison between the behavioristic and cognitive approaches to the question of learning. The behaviorist has argued that such reports are another verbal response which is irrelevant to the conditioning phenomenon, while the cognitive psychologist has viewed these as indicative of awareness, which is the necessary mediating process between situational stimuli and resulting behavior.

While the legitimacy and importance of post-experimental accounts of subject awareness is at the heart of this controversy, it has rarely been linked to the recent research concerning subject-role behavior. This is unfortunate, since a basic, but untested, assumption of the cognitive approach is that the subject is motivated to cooperate with the experimenter. In fact, the suggestion that awareness produces conditioning is based on the legitimacy of the good-subject role.
Perhaps the most detailed and extended attention to this question has been within the verbal operant-conditioning paradigms of Greenspoon (1955) and Taffel (1955) in which a verbal response class is systematically reinforced by the experimenter with the words "good" or "hmmhmm." These studies concluded that the increase of emitted reinforced responses took place without the subject's being aware of the reinforcement contingency, since subjects who conditioned did not seem to be able to describe the correct contingency. Since then a great deal of controversy has resulted over the role of awareness in producing the conditioning effect.

Probably the most elaborate line of theorizing and methodological development within the cognitive approach has been by Dulaney (1962, 1968) whose work represents the extreme cognitive position. His Theory of Propositional Control specifies that, in general, a person behaves as a function of his willfull compliance with rules which he deduces about his environment. Fishbein has extended the conceptualization to the study of attitudes, intentions, and behavior, and recently others have extended the model to voting behavior (Fishbein & Coombs, 1971; Williams, Weber, Haaland, Mueller, & Craig, 1973). Since its original conception was in terms of subject behavior in the verbal-conditioning laboratory, the present research is quite relevant in that it offers the possibility of extending Dulaney's theory to a closely-related kind of experimental situation, the classical conditioning of meaning experiment. In the present study,
empirical support of Dulaney's theoretical model would offer a further insight into the ways in which the manipulated cues affect responding, since validation of it would suggest that subjects respond to them in a very rational fashion.

According to the Theory of Propositional Control, what a subject does in an experiment is a function of his conscious cooperation with the propositions he has discerned about the situation. That is, his behavior depends upon his knowledge about the experiment (the rules) and his feelings about it (the ascriptors). A subject is only able to give the conditioned response if he knows the specific reinforcement contingency, and this knowledge is critical, since the process begins here, and without contingency awareness, other aspects of awareness would not affect behavior. Once the subject becomes aware of the contingency rule, his subjective evaluation of the reinforcement becomes important in influencing his understanding of the experimenter's behavioral hypothesis, for if the reinforcement is aversive, the subject is unlikely to infer that he is supposed to behave in order to maximize it. Thus the behavioral hypothesis is the next component of awareness which is produced by the multiplicative relationship of the contingency rule and the subjective value of the reinforcement. If either of these is zero, the behavioral hypothesis will be zero. If either is negative, the subject will assume he is supposed to emit a response which is actually counter to the experimenter's expectations.
Once the behavioral hypothesis is established, the subject's motivation to comply with the experimenter becomes important, since this must have a positive value in order for the subject to condition according to expectations. Thus behavioral hypothesis and motivation-to-comply multiply to produce behavioral intention, which because of the volitional nature of such behavior, is the best predictor of performance on the dependent variable. Thus, the formal theoretical proposition is indicated by the following model:

\[
\begin{align*}
\text{(Awareness of Contingency Rule)} & \times \text{(Subjective Value of Reinforcement)} \\
\downarrow & \\
\text{(Behavioral Hypothesis)} & \times \text{(Motivation to Comply)} \\
\downarrow & \\
\text{(Behavioral Intention)} & \\
\downarrow & \\
\text{("Conditioning")}
\end{align*}
\]

where the arrows read as "influences." Each of these components measured through a post-experimental questionnaire progresses from the subject's knowledge of the contingencies to his inference regarding the experimenter's hypothesis, to his behavioral intention, to the actual behavior. The first and second steps are influenced by the ascriptor values indicating how much the subject desires the reinforcement and how much he wants to cooperate with the hypothesis. Thus the first requirement for learning is contingency awareness, which
(unless the subjective value of the reinforcement is negative) produces a knowledge of the hypothesis. In the standard conditioning situation, contingency and hypothesis awareness are highly correlated. The effects of contingency awareness on conditioning are mediated through the production of the hypothesis, so that if hypothesis awareness is partialled out, the relationship between contingency awareness and conditioning falls to zero. Similarly, the effects of hypothesis awareness depend upon the production of a behavioral intention, so that partialling out the behavioral intention leaves the relationship between the behavioral hypothesis and conditioning at zero. The dependent variable is most highly correlated with behavioral intention, next with behavioral hypothesis, next with contingency awareness.

This formulation is useful for approaching the study of awareness in the present experiment, since the assumptions are compatible with those of the present research: that in order to cooperate with the demand characteristics of an experiment, two information requirements must be satisfied. The first is hypothesis awareness, or knowledge of what is expected, and the second is the absence of a reason for not cooperating. That is, without explicit reason, subjects will be good subjects and cooperate, but when given reasons not to cooperate (i.e., evaluation apprehension), they will resist behaving in line with hypothesis cues.
Empirical Support for the Dulaney Model

Previous research by Dulaney suggests that his model is very useful for describing behavior in several variations of the verbal operant conditioning experiment. For example, the components have been found to account for over 90 percent of the variance in behavioral intention and conditioning which both correlate highly with each other (Dulaney, 1962, 1968). Similarly, partialling out intervening terms reduces the correlation of distal terms to zero. Thus, with behavioral intention out, the correlation between behavioral hypothesis and conditioning falls to zero. Likewise, correlating conditioning with each of the terms without partialling produces an ordered model in that the best predictor of conditioning is behavioral intention, the next is behavioral hypothesis, and the weakest is contingency awareness. In a converging demonstration of the adequacy of the model, subjects with no contingency awareness are compared with fully-aware subjects and predictions are shown to hold only in the later case. When contingency awareness was absent, the relationship between the resulting behavior and each of the other terms dropped to zero. All of these findings have been replicated by others (Bottom, 1972; Doctor, 1971; Uleman, 1971a, 1971b; Uleman & Vandenbox, 1971). Thus, there are converging lines of evidence based on many different analyses from several different laboratories which suggest that the effects of these components work in a systematic fashion such that subject motivation is dependent first upon
hypothesis awareness, and then upon cooperation.

In the present experiment it is possible to adapt the Dulaney formulation through some translations and modifications. Contingency awareness (CA) is analogous and can be measured by asking subjects which nonsense syllables were paired with pleasant and unpleasant words (i.e., Insko & Oakes, 1966; Page, 1969). The behavioral hypothesis in Dulaney's paradigm is essentially the same as Page's demand awareness (DA) or hypothesis awareness (HA) which is measured by asking subjects how they think the experimenter expected them to rate various key syllables. The measure of motivation to comply (MC) is identical (i.e., "How much did you want to behave according to the way the experimenter expected you to"). Since past research has revealed some problems with this component (Dulaney, personal communication, 1973; Bottom, 1972; Uleman, 1972, b) an item from Page's (1969) post-experimental questionnaire could also be used as perhaps a more successful operationalization.

Some of the terms do not translate so readily, however. Behavioral intention produces the most serious problem, since it becomes meaningless in the present case where the dependent variable involves a single response rather than a pattern of responses over time, as in the verbal operant paradigm. Asking a subject how he intended to rate a nonsense syllable is likely to be translated by the subject as "How did you rate the syllable" since it is difficult to recognize any possible discrepancy between the intention to perform
this specific response and the actual performance of it. Subjective value of the reinforcement is also a less meaningful term in the classical conditioning experiment, since the evaluative dimensions of the words used as UCSs are much more variable than the single reinforcement given in the verbal operant paradigm.

In spite of the loss of these components in the model, the basic rationales is still testable. Conditioning should be a multiplicate function of HA and MC. And CA and evaluation awareness (EA), which are types of awareness that are being manipulated in this study, should affect conditioning through their operation on HA and MC respectively. Thus, a theoretical network appropriate for the present study would seem to be

\[
\begin{align*}
&\text{CA} \\
&\downarrow \\
&\text{HA} \times \text{MC} \rightarrow \text{DV} \\
&\uparrow \\
&\text{EA}
\end{align*}
\]

where the arrows represent one-way causal paths. This model is then testable through the following predictions:

1) DV should be more highly correlated with HA x MC than with EA or CA
2) Partialling HA x MC reduces the relationship of DV to EA and CA to zero.
Support for these predictions would suggest that subjects select their responses in a highly cognitive, conscious manner consistent with the formulation of propositional control. Failure to support these predictions would suggest a view of subject performance more in keeping with the suggestions of Orne and Rosenberg that subjects have little conscious control over these behaviors.

According to Orne, the subject may not employ a conscious motivation to comply even though he systematically does cooperate, and so one would not expect the MC term to have much predictive value on the dependent variable, nor to aid in prediction when it is multiplied by hypothesis awareness. To the extent that it could be shown that the MC measure was valid, and yet did not correlate with DV, with or without being multiplied by HA, Orne's position would receive empirical support.

Rosenberg would take a similar stand with regard to the measurement of EA, since he has previously suggested that subjects do not recognize the extent to which they are responding according to evaluative concerns, and so would not be expected to insightfully report them. Thus, to the extent that it could be shown that the EA measure was valid, and yet did not correlate with DV, or did not contribute to DV through the MC term, Rosenberg's position would be supported.
Overview of the Research Objectives

As this chapter indicates, there are several issues addressed by the research, and it should be noted that the study was designed to investigate subject behavior at several levels of observation.

The first and most important question regards the effects of situational manipulations of types of demand characteristics on subject behavior. By independently varying hypothesis and evaluation cues, incidence of cooperation and resistance could be observed. By choosing the classical conditioning of attitudes paradigm, these manipulations could be made orthogonal. Since previous pilot research indicated that credibility problems arose from informing subjects of the experimental hypothesis directly, hypothesis awareness was manipulated by enhancing contingency awareness. Thus contingency information served as the first variable, which was crossed with evaluation cues in the form of the presence or absence of a conformity test and the presence or absence of an accomplice who suggested that the experiment was about conformity. This allowed for the operational specification of the good and apprehensive roles, which were linked directly with experimental procedures.

In addition to the situational manipulations, a personality variable provided further validation of the major predictions by representing a chronic manipulation (McGuire, 1968) of the same construct being tapped by the acute manipulation of evaluation cues, i.e., the probability of evaluation
apprehensiveness. Thus the SD scale of Crowne and Marlow was dichotomized as an additional independent variable, and predictions regarding its interaction with the cues were tested.

The need for valid manipulation checks of the situational manipulations gave rise to an additional set of research objectives: to trace the effects of these cues on post-experimental awareness. By performing convergent and discriminate validation checks, the validity of the cues could be assessed in the most thorough manner to date. Thus, the effects of contingency information on contingency and hypothesis awareness could be assessed, and the effects of evaluation cues on evaluation awareness could be assessed to provide evidence of convergent validity. The effects of contingency information on evaluation awareness and the effects of the evaluation cues on contingency and hypothesis awareness could be assessed to provide evidence of discriminant validity. Thus, in combination with the cue manipulations, these measures provided an opportunity for a conceptualization and operationalization of the cooperative and apprehensive roles which is perhaps the clearest in the murky history of these concepts.

Finally, the need to gather awareness data allowed the opportunity to investigate a secondary question regarding the degree to which such awareness reports affect behavior according to the model of propositional control. To the degree that predictions derived from the Dulaney model held,
the effects of demand characteristics could be further specified as being the result of a very conscious decision-making process.

Thus, the major questions which were asked involved 1) the effects of hypothesis and evaluation cues, as well as the subjects need for social approval, on behavior, 2) the relationships of such manipulations to subject reports of awareness, and 3) the relationship between various components of awareness and their relationship to overt behavior. In this fashion, there was a triangulation of observations which is diagrammed in Figure 1.
RELATIONSHIPS AMONG QUESTIONS BEING ASKED

TYPE I
ESSENTIAL VALIDITY CHECKS

MANIPULATION OF DEMAND CHARACTERISTICS
(HYPOTHESIS AND EVALUATION CUES)

SUBJECT AWARENESS REPORTS ON POST-EXPERIMENTAL QUESTIONNAIRE

TYPE II
PRIMARY QUESTION:
DETERMINANTS OF SUBJECT ROLES

TYPE III
SECONDARY QUESTION:
ROLE OF AWARENESS IN CONDITIONING PARADIGM

SUBJECT COOPERATION ON CONDITIONING TASK

Fig. 1. Pictorial representation of research objectives.
II. METHOD

Design

The experimental design employed three situational manipulations and one individual difference factor to produce a $2 \times 2 \times 2 \times 2$ factorial design. The four factors were contingency information (hypothesis cue), test and accomplice (evaluation cues), and high or low social desirability. Operationalization of each of these independent variables is described below.

Factor 1: Contingency Information (CI)

Since previous pilot research indicated that subjects become very suspicious of explicit information about the experimenter's hypothesis, hypothesis awareness was manipulated through the delivery of contingency information. Subjects in the Information (I) conditions were given explicit information regarding which syllables would be paired with positive and negative words, while subjects in the No-Information (NI) conditions were not given such information. The exact wording and procedures of this manipulation, as well as those for the next two factors, are described in the procedure section which follows.
Factor 2: Test (T)

The second factor served as one of the evaluation cue factors, in which a personality test obviously dealing with conformity either was or was not administered to subjects at the beginning of the experimental session.

Factor 3: Accomplice (A)

The third factor and second evaluation cue factor was adapted from Rosnow, Goodstadt, Suls and Gitter (1973) and consisted of an accomplice who either did or did not ask a question in the middle of the experiment about the purpose of the study, suggesting that it dealt with conformity.

Factor 4: Social Desirability

The fourth factor consisted of two levels of scores on the Marlow-Crowne Social Desirability Scale (SD) which was administered some time before the experimental sessions were run. Scores on this scale were split at the median, creating the dichotomized variable.

Subjects

Subjects were University of New Hampshire Introductory Psychology students who were required to participate in a number of psychology experiments as part of their laboratory experience. They were recruited from two large lecture sections because of the necessity of having them take the Social Desirability scale before the experiment was run.
Procedure

Collection of Individual Difference Measure

Approximately four weeks before the experiment was run, subjects were given the Marlow-Crowne Social Desirability Scale which is presented in Appendix B. This measure was collected while the subjects were in class, by their instructors, who explained that it was part of a class demonstration planned for a later date. The experimenter was not present. Responses were made on IBM scanning sheets and machine scored.

Experimental Procedure

Subjects signed up to participate in experimental sessions depending upon their score on the SD measure. Each name listed on class rosters was assigned a letter depending upon whether the individual scored above or below the median. A third letter was given to persons who did not take the measure. Subjects then signed up for experimental sessions within the appropriate letter slot so that each session had approximately the same number of high scorers, low scorers, and no scorers. Subjects were not aware of the relationship between their letter designation and their scores on the SD scale. Experimental sessions lasted for approximately one hour, and were conducted within a two and one-half week period, in similar numbers of morning, afternoon, and evening sessions. Experimental conditions were varied randomly, according to a blind procedure described below.
After subjects were checked in, they were led to a group testing room in which 16 chairs faced a screen, and a projector and tape recorder were located in the back where the experimenter could operate them. Subjects were given notebooks and pencils, were told that the experiment consisted of several parts, each of which had tape-recorded instructions. The rest of the experimental instructions were then delivered by tape.

Several versions of the same master tape were recorded so that instructional manipulations could be varied while the rest of the information and its delivery remained constant.

Standard Staats and Staats Procedure

The experimental procedure which served as a core and to which the various cue manipulations were added was followed for subjects in all conditions. This occurred in the following manner:

1. The tape recorder was turned on and subjects heard the General Instructions (see Appendix A) which were fashioned after Staats and Staats (1957), describing the study as one of intermodality learning, and requesting that subjects pay attention to the nonsense syllables to be flashed on the screen in front of them, and at the same time to repeat the word on the tape recorder in a loud, clear voice, together as a group. After the tape asked for any questions, and the experimenter walked to the front of the room to check if there were any, the lights were turned off, and the tape recorder and slide projector turned on simultaneously, so that after
each word was read from the tape, the experimenter advanced the projector to the next slide. In this way, all subjects heard the same auditory stimuli in conjunction with the nonsense syllables. The only qualification of this involves the control procedure in which the key syllables were switched every other session, so that half of the time YOF was paired with positive words and WUH with negative, while the other half of the time YOF was paired with negative and WUH with positive. This insured that any pre-experimental attitudes about the syllables (particularly to WUH, which is very similar to the university's initials UNH) could not confound the dependent variable.

2. After all the syllables had been shown, the lights were turned on again and the Dependent Variable Rating Scale (see Appendix C) was passed out to each subject. The tape then delivered the Dependent Variable Instructions (see Appendix A). This information instructed subjects to mark each syllable in the booklet according to how pleasant or unpleasant they felt it was, and also the check whether the syllable was presented on the screen or not. The booklet contained 16 syllables, 5 of which had actually been presented. At the completion of these instructions, the tape asked for any questions which the subject might have and the experimenter walked to the front of the room to see if there were any.

3. After all subjects finished the rating scale, the scales were collected, and the post-experimental questionnaire
(see Appendix E) was passed out, and instructions for it presented on the tape (see Appendix A). These instructions emphasized that the questionnaire was the most important part of the study and that if there were any problems in filling it out, subjects were urged to ask the experimenter.

4. After all subjects finished the post-experimental questionnaire, the questionnaires were collected and the experimenter urged the participants not to speak of the experiment with their friends. Subjects were then dismissed. Debriefing was conducted several weeks later in the introductory classes, and the major analysis of the independent variables was reported. In line with the department's policy concerning the use of introductory students as subjects, each subject wrote a debriefing resume which listed the major area of research, major question being asked, independent and dependent variables, results, implications, and subjective evaluation. These reports were read by the experimenter and credit was given to the subject for participating if the majority of the information was correct (criteria focused on the correct specification of the independent and dependent variables).

Manipulations of Basic Procedure

Subjects in the No Information-No Test-No Accomplice condition experienced the procedures described above. The other conditions were created by the addition of various cues to this basic procedure in the following manner:
Contingency Information. During the General Instructions described in step 1, a sentence was dubbed into the tape which described the contingencies (see Appendix A for exact placement). This sentence indicated that during the pairing of the syllables and words, "it will become obvious that two of the syllables are special. They are special because one of them, YOF will always be paired with words of positive or pleasant meaning, and WUH will always be paired with words of neutral meaning." (In line with the control for pre-experimental attitudes regarding the syllables, this sentence was changed half of the time by dubbing YOF for WUH, and vice versa, so that the contingencies were switched while the rest of the sentence remained constant.) Subjects in the No Information conditions did not hear this spliced-in sentence.

Test. The Test manipulation was made at the beginning of the experiment, just after subjects were told that the experimental instructions would be delivered by tape and before step 1 of the standard procedure described above. In this Test condition (T) a personality test was passed out to each subject, and the first instructions heard on the tape described instructions for completing it (see Appendix A). The items on the test were statements which dealt with the conformity process (see Appendix D) such as "I have had some people tell me that I am too gullible," "I think its important not to 'rock the boat,'" and "I think cooperation is more important than individualism." These items were written by the author in an attempt to make the issue of conformity a salient one.
and suggest that the test was measuring a person's tendency to conform. Many of the items were fashioned after the MMPI scale, which was the source of a similar manipulation by Page (1971).

After subjects completed the scale, they were asked to put their names and introductory class section numbers on it, in an attempt to enhance evaluation apprehension.

Accomplice. In the Accomplice conditions, an accomplice of the experimenter asked a question directly after the instructions for the dependent variable described in step 2 of the standard procedures, and before subjects began rating the syllables. The confederate was present at every session, and according to a schedule to which the experimenter was blind, either raised his hand to ask or did not raise his hand to ask the following question:

A friend of mine was in this experiment last week. Isn't this the experiment which is really a test of conformity, rather than intermodality learning? Aren't you really deceiving us, and actually showing that those who rate the nonsense syllables according to the way they were paired are really conforming or gullible?"

The experimenter's reaction to this was an extended stare (5 to 10 seconds) with the comment, "I'll answer that question when the experiment is over, if you don't mind."

Scoring Procedures

There were seven variables to be scored, and they are described in the following order: 1) Individual Difference Measure, 2) Dependent Variable, and 3) Mediators and Manipulation Checks.
Individual Difference Measure. The scoring of the Social Desirability Scale, 33 items in length, followed the procedures set forth by Crowne and Marlow (1964). Direction of wording was counterbalanced in the measure itself. The scores on the SD scale could range between 0 and 33, with a high score indicating high social desirability. The actual distribution observed is presented in Figure 2, where it can be seen that the range varied from 1 to 29 with a standard deviation of 5.321. The median was 14.50, and so subjects who scored 14 and below were assigned to the Low SD condition, while subjects who scored 15 and above were assigned to the High SD condition.

Dependent Variable (DV). The dependent variable, that of amount of conditioning, was scored in the manner of previous studies (Page, 1969; Weber & Riddell, 1973). Specifically, numerical values of 1 to 7 were assigned to responses on the pleasant-unpleasant dimension. The score on the positively-paired syllable was then subtracted from the score on the negatively-paired syllable yielding a measure of conditioning which ranged between 6 (indicating full conditioning) to -6 (indicating conditioning in the opposite direction), with zero indicating no conditioning effect at all.

Since the contingencies of the key syllables were switched every other session, this scoring procedure insured that any pre-experimental attitudes toward the specific syllables could not confound the results.
Fig. 2. Distribution of SD scores with median split indicated.
Mediators and Manipulation Checks. 1) Contingency Awareness (CA). Contingency awareness was measured by items 6 and 7 on the post-experimental questionnaire (see Appendix E), designed after Page (1969). However, this measure was scored on a continuum rather than on a dichotomy so that the maximum amount of information could be used in the analysis of awareness. The item asked the subject which syllable was paired with positive and negative words. The correct answer to these questions, two CVC nonsense syllables, involved six letters, three for each question, and one point was given for each correct letter identified. Thus, the CA measure ranged between zero and 6, with 6 indicating full awareness, and zero indicating no awareness. In the few cases (5/385) in which subjects indicated the positively paired CVC in response to the question regarding the negatively-paired syllable (and vice versa), a score of -6 was assigned, indicating awareness of the opposite contingency.

2) Hypothesis Awareness (HA). Hypothesis awareness was measured by items 8 and 9 on the post-experimental questionnaire (see Appendix E). Each dimension was scaled from 1 to 7, and similar to the manner in which the DV was scored, the score on the positively-paired syllable was subtracted from the score on the negatively-paired syllable, yielding a measure of HA which ranged between 6 (indicating full awareness) to zero (indicating no awareness) to -6 (indicating awareness of the opposite expectation).
3) Evaluation Awareness (EA). Three items were used to measure evaluation awareness. The first was item 10 which was fashioned after Dulaney's method of measuring awareness in a multiple-choice format. Subjects who checked alternative A were assigned a score of 1, B a score of 2, C a score of 3, D a score of 4, E a score of 5, and F a score of 6.

The second measure of EA consisted of items 11 and 12, which were short-answer essays asking the subject "If a person rated the nonsense syllables (or in 12, did not rate the nonsense syllables) according to the pleasant or unpleasant words they were paired with, what would this indicate about his personality or intelligence?"

Answers to this item required more elaborate scoring procedures. If a subject indicated a strong statement of some negative quality in 11 and some positive quality in 12, (such as "He's gullible" and "He's got a mind of his own") then he was assigned a score of -2. A milder statement, indicating less certainty (i.e., "I'm not sure but . . .") received a score of -1. Similarly, if the subject indicated a strong statement of some positive quality in 11 and some negative quality in 12 (such as "He's intelligent" versus "He didn't pay attention") a score of +2 was assigned. A milder, less certain statement in the same direction was given a -1. A zero was assigned if answers to both items were both positive or both negative (thereby cancelling each other) or if the subject stated that he didn't know, or if he merely restated the question.
Since this item entailed a subjective judgment, a reliability assessment was performed by having another judge, blind to the evaluation cue conditions and the purpose of the experiment, rate the answers according to the above system. Since the interjudge reliability correlation was quite high (.9357), the first judge's ratings were used as the score.

The third measure of evaluation awareness involved answers to item 13, which asked the subject to rate a person who would condition on a series of dimensions. The gullible versus not gullible dimension and the individualistic versus conforming dimension were used, and responses on these scales were scored from 1 to 7.

4) Motivation to Comply (MC). Two items were used to measure motivation to comply. The first, item 14 was taken directly from the Dulaney paradigm (Dulaney, 1968; personal communication, 1973) and simply asked the subject "During the experiment, did you want to rate the syllables the way you thought the experimenter expected you to?" Answers to this were scored along a 1-to-7 continuum. The second item (15) was fashioned after Page (1971) which served as an elaboration of the first item, involving the subject's desire to help the hypothesis along. This was also scored along a 1-to-7 continuum. Both items were scored in a negative direction with 1 indicating much desire to comply, and 7 indicating no desire, or desire to the opposite. These items yielded measures of MC which were sublabeled MC_E (item 14, for the experimenter's expectations) and MC_H (item 15, for help
Predictions

Type I: Manipulation Checks

Convergent Validity Checks.

1. Contingency Information Factor (C.I.)
   a. A significant effect of C.I. was predicted such that more Contingency Awareness as measured by items 6 and 7 would occur at I than at NI. This was tested by a single factor analysis of variance.
   b. A significant effect of C.I. was predicted such that more Hypothesis Awareness, as measured by items 8 and 9 would occur at I than at NI. This was tested by a single factor analysis of variance.

2. Evaluation Cue Factors. This was tested by a single factor analysis of variance. It was predicted that more evaluation awareness would be measured by items 10, 11, 12, and 13 when more cues were present than when less cues were present. Since previous pilot work with the dependent variable suggested stronger effects for the accomplice than for the test manipulation, a pattern of means was predicted such than an interaction between the two factors would be yielded in the following manner:
Discriminant Validity Checks.

1. Contingency Information Factor. It was predicted that there would be no significant effect of CI on various measures of EA (items 10, 11 and 12, 13), as tested by a single factor analysis of variance.

2. Evaluation Cue Factors.
   a. It was predicted that there would be no significant effect of the evaluation cues on CA.
   b. It was predicted that there would be no significant effect of the evaluation cues on HA, as tested by a single factor analysis of variance.

Type II: The Effects of Manipulations on Conditioning

A four-factor interaction between the two evaluation cues, contingency information, and social desirability was predicted which is represented in Figure 3. The nature of this interaction may be briefly described as follows: that the relative effects of the evaluation cues in decreasing conditioning would depend upon contingency information and subject
Fig. 3. Predicted interaction of independent variables on conditioning based on pilot work.
need for approval such that the evaluation cues would be more powerful when contingency information was present and among High SD subjects, than that which would be predicted on the basis of the CI and SD factors alone.

With regard to specific factors, the interaction can be specified as it relates to the following hypotheses, listed in order of importance:

H1. Subjects will cooperate unless there are clear evalu­ative reasons for not cooperating, indicating evidence of both the good and apprehensive roles.

Thus, it was predicted that the significant simple effects of contingency information in increasing conditioning would be most numerous when no evaluation cues were present, least numerous when both evaluation cues were present, and of intermediate number when only one evaluation cue was present.

Since this relationship was expected to be stronger for High SD than for Low SD subjects, any non-significant differences in the one or both cue conditions were expected to occur among Low SD subjects, and any significant differences in the no-cue condition were expected to occur among High SD subjects.

H2. The effectiveness of the evaluation cues in decreasing conditioning depends upon the effective communication of hypothesis cues.

Thus, it was predicted that the significant simple effects of evaluation cues in decreasing conditioning would
be more numerous at I conditions than at NI conditions.

Since this relationship was expected to be stronger for the High SD than the Low SD subjects, any non-significant differences in the I conditions were expected to occur among Low SD subjects (and/or when the other evaluation cue was absent) and any significant differences which occurred in the NI conditions were expected to occur among High SD subjects (and/or when the other evaluation cue was present).

H3. Subjects' needs for social approval interacts with hypothesis and evaluation cues such that the SD factor is more effective when many cues are present.

Thus it was predicted that there would be the greatest number of simple effects of SD as the number of cues increased, and that these effects would occur such that, relative to Low SD subjects, High SD subjects would condition more when no evaluation cues were present, especially when contingency information was given, but condition less when evaluation cues were present, especially when contingency information was given.

Since previous pilot work suggested an interaction among the evaluation cues, it was further predicted that the effectiveness of the test would be dependent upon the presence of the accomplice, but that effectiveness of the accomplice would not depend upon the presence of the test.

This prediction rests on the assumption that the accomplice produces more evaluation awareness than does the
test, which could be checked in the current study with the various measures of evaluation awareness. Should the measures of evaluation awareness suggest they are additive, this prediction would be refined accordingly.

Type III: The Study of Awareness

H4. The effects of evaluation awareness depends upon the prior presence of hypothesis awareness. This hypothesis is tested by the following predictions:

1. When HA = 1 or 2, correlation between EA and DV = 0.
2. When HA = 6 or 7, correlation between EA and DV is positive and significant.
3. When EA = 0, correlation between HA and DV is still positive and significant.

H5. The effects of awareness operate according to the Dulaney theory of propositional control. The model to be tested within the current paradigm was

\[ \text{CA} \]
\[ \downarrow \]
\[ \text{HA} \times \text{MC} \longrightarrow \text{DV} \]
\[ \uparrow \]
\[ \text{EA} \]

where the arrows represent one-way causal paths. The multiplicative relationship between HA and MC indicates that, when either term is zero, the resulting conditioning is also zero.
The following predictions test this model:

H5a. A "ladder" of correlations should appear between the components such that conditioning (DV) correlates from most to least with

- HA X MC
- HA
- CA

H5b. The partialling of intervening terms should reduce the correlation between the adjacent terms to zero. Thus, the correlation between CA and DV with HA x MC partialled = 0,
the correlation between EA and DV with HA x MC partialled = 0.
III. RESULTS

Results are reported within three major sections: manipulation checks (Part I); the effects of manipulations on behavior (Part II) and correlational data from the post-experimental questionnaire relating behavior to awareness (Part III). Refinements of some predictions in Part II are based on findings in Part I.

A total of 385 subjects participated in the experimental session, but only 322 of these had previously taken the SD measure. Since care was taken to insure that subjects filling out the SD scale did not associate that exercise with the experiment, the possibility that the 63 subjects who did not fill out the scale being different from the 322 who did seemed minimal, and the total available data were used in each analysis. Therefore, the total N for the analyses reported in Part I and Part III was 385, whereas in Part II, in which SD was a factor, it was 322.

Part I: Manipulation Checks

The convergent validity data present a reasonably clear pattern which suggests that each of the independent situational manipulations were successful in producing the desired effects on subject awareness. Most measures of discriminant validity show that these manipulations were not confounded. The results of the convergent validity checks are presented first.
Convergent Validity

**Contingency Information Factor.** The delivery of contingency information produced the predicted effects on both contingency and hypothesis awareness. Subjects in the I condition had a mean CA score of 5.51, while those in the NI condition had a mean of 3.88. This difference was significant beyond the .001 level ($F = 55.68, df 1/380$).*

The effects of the information were similar on the measure of hypothesis awareness. Those in the I condition had a mean HA of 5.41, while those in the NI condition had a mean of 3.67. This difference was also significant beyond the .001 level ($F = 49.11, df 1/383$).

**Evaluation Cues.** There were four measures of evaluation awareness used to assess the validity of the evaluation cue factors.

The first item was 10, fashioned after Dulaney's measure of awareness in multiple-choice format. The means on this item are presented in Table 1, along with the $2 \times 2$ ANOVA which indicated that the Test factor was significant ($p < .05$), while the Accomplice factor was not. Contrary to predictions, there was no significant interaction between the two factors.

The second item was the essay items (11 and 12). The means obtained for this are shown in Table 2, along with the corresponding analysis of variance. As can be seen, evaluation

*All ANOVAS were performed with unweighted means analysis.*
TABLE 1

EFFECT OF EVALUATION CUES ON DULANEY-TYPE MULTIPLE CHOICE ITEM OF EVALUATION AWARENESS

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p &lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>8.077</td>
<td>1</td>
<td>8.077</td>
<td>4.599</td>
<td>.05</td>
</tr>
<tr>
<td>Accom</td>
<td>.474</td>
<td>1</td>
<td>.474</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T X A</td>
<td>.775</td>
<td>1</td>
<td>.775</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>668.9</td>
<td>381</td>
<td>1.756</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Means

<table>
<thead>
<tr>
<th></th>
<th>Test</th>
<th>No Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accom</td>
<td>3.37</td>
<td>2.99</td>
</tr>
<tr>
<td>No Accom</td>
<td>3.21</td>
<td>3.01</td>
</tr>
</tbody>
</table>
TABLE 2

EFFECT OF EVALUATION CUES ON ESSAY ITEM OF EVALUATION AWARENESS

Means

<table>
<thead>
<tr>
<th></th>
<th>Test</th>
<th>No Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accom</td>
<td>-.500</td>
<td>.200</td>
</tr>
<tr>
<td>No Accom</td>
<td>-.175</td>
<td>.474</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p &lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>43.675</td>
<td>1</td>
<td>43.675</td>
<td>19.089</td>
<td>.01</td>
</tr>
<tr>
<td>Accom</td>
<td>8.610</td>
<td>1</td>
<td>8.610</td>
<td>3.763</td>
<td>.05</td>
</tr>
<tr>
<td>T X A</td>
<td>.063</td>
<td>1</td>
<td>.063</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>869.721</td>
<td>380</td>
<td>2.288</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
cues produced significant effects, the Test factor reaching significance at the .01 level, the Accomplice reaching it at the .05. Again, the interaction between the two factors was not significant.

The two additional measures of evaluation awareness asked subjects to attribute degrees of gullibility and conformity to a subject who conditioned. The means obtained from these items are presented in Tables 3 and 4, where it can be seen that the Test factor was significantly related to attribute conformity ($p < .05$) but only approached significance for gullibility ($p < .10$). In both cases the Accomplice factor was not significant, nor was the interaction of it with the Test factor.

Overall, the convergent validity checks suggest that the manipulations produced the desired effects on awareness. The contingency information is especially clear-cut. The evaluation cues show a more complex pattern. While the essay item showed significant effects of both of the cues, the other measures indicated the effects of the Test manipulation only, and no measures yielded any significant interactions between the cues. Since the essay item was more sensitive to the Accomplice as well as to the Test manipulation, it seemed that it was the best measure of evaluation awareness.

**Discriminant Validity**

The data concerning discriminant validation allowed for the deduction concerning the orthogonality of the
TABLE 3

EFFECT OF EVALUATION CUES ON ATTRIBUTED GULLIBILITY

Means

<table>
<thead>
<tr>
<th></th>
<th>Test</th>
<th>No Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accom</td>
<td>4.635</td>
<td>4.357</td>
</tr>
<tr>
<td>No Accom</td>
<td>4.451</td>
<td>4.216</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p &lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>6.315</td>
<td>1</td>
<td>6.315</td>
<td>3.785</td>
<td>.10</td>
</tr>
<tr>
<td>Accom</td>
<td>2.536</td>
<td>1</td>
<td>2.536</td>
<td>1.520</td>
<td></td>
</tr>
<tr>
<td>T X A</td>
<td>.046</td>
<td>1</td>
<td>.046</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>634.001</td>
<td>380</td>
<td>1.668</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 4

EFFECT OF EVALUATION CUES ON ATTRIBUTED CONFORMITY

Means

<table>
<thead>
<tr>
<th></th>
<th>Test</th>
<th>No Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accom</td>
<td>4.950</td>
<td>4.947</td>
</tr>
<tr>
<td>No Accom</td>
<td>4.793</td>
<td>4.319</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>14.789</td>
<td>1</td>
<td>14.789</td>
<td>6.462</td>
<td>.05</td>
</tr>
<tr>
<td>Accom</td>
<td>5.461</td>
<td>1</td>
<td>5.461</td>
<td>2.386</td>
<td></td>
</tr>
<tr>
<td>TXA</td>
<td>5.324</td>
<td>1</td>
<td>5.324</td>
<td>2.231</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>869.701</td>
<td>380</td>
<td>2.287</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
manipulations, a major aim of this research. By checking the impact of each of the factors on the corresponding measure of awareness for the other, it was possible to see if the manipulations produced confounding.

Contingency Information. The impact of contingency information on the measures of evaluation awareness produced a complicated pattern. On the first measure (10) the Information group showed slightly less awareness (3.02) than did the No Information group (3.26). This difference was significant beyond the .05 level. Since the effect of information was to enhance conditioning (results to be reported shortly), this result was particularly troublesome, since it suggested a possible confounding between evaluation awareness and contingency awareness. However, the correlation between this measure and conditioning was only -.0477, which is not significant. The ANOVA table for this measure of discriminant validation is reported in Table 5. The implications of this significant finding will be elaborated at the end of this section.

The second measure of evaluation awareness, the essay item, which produced clear relationships to the evaluation cues, bore no relationship to the Information manipulation. Here the means were not significantly different from each other ($F = .061$, df 1/383).

The final two measures of evaluation awareness did show significant relationships to the Information factor, as shown by Table 6. On the Individualistic-Conforming dimension,
TABLE 5

EFFECT OF CONTINGENCY INFORMATION ON DULANEY-TYPE ITEM OF EVALUATION AWARENESS

<table>
<thead>
<tr>
<th>Item</th>
<th>N</th>
<th>Mean</th>
<th>Contingency Info</th>
<th>No Contingency Info</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>182</td>
<td>3.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>203</td>
<td>3.26</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p  &lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI</td>
<td>1</td>
<td>5.97</td>
<td>5.97</td>
<td>3.40</td>
<td>.05</td>
</tr>
<tr>
<td>Error</td>
<td>383</td>
<td>672.60</td>
<td>1.756</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 6

EFFECT OF CONTINGENCY INFORMATION ON ATTRIBUTED GULLIBILITY AND CONFORMITY

I. On Attributed Conformity

<table>
<thead>
<tr>
<th>Means</th>
<th>I</th>
<th>NI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.97</td>
<td>4.45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p &lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI</td>
<td>25.39</td>
<td>1</td>
<td>25.39</td>
<td>11.24</td>
<td>.001</td>
</tr>
<tr>
<td>Error</td>
<td>865.30</td>
<td>383</td>
<td>2.259</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

II. On Attributed Gullibility

<table>
<thead>
<tr>
<th>Means</th>
<th>I</th>
<th>NI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.66</td>
<td>4.30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p &lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI</td>
<td>612.36</td>
<td>1</td>
<td>12.36</td>
<td>7.45</td>
<td>.001</td>
</tr>
<tr>
<td>Error</td>
<td>635.64</td>
<td>383</td>
<td>1.66</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the Information group had a mean of 4.97, showing more perception of the conformity attribute than did the No Information group, which had a mean of 4.45. This difference was significant beyond the .001 level ($F = 11.24$, $df = 1/383$).

On the Gullibility dimension, the Information manipulation also showed an effect. Here the condition scored a mean of 4.66, whereas the NI scored 4.30. This suggests that the Information group saw the conditioning as more associated with gullibility than did the NI group. This effect was significant also beyond the .001 level ($F = 7.45$, $df = 1/383$). The implications for this finding will also be reviewed at the end of this section.

**Evaluation Cues.** The discriminant validation data for the evaluation cues were much more clear-cut, where neither of the cues showed significant relationships to the measures of contingency or hypothesis awareness. On both measures, the $2 \times 2$ analysis of variance yielded non-significant effects of the cues, as well as non-significant interactions.

**Summary of Validity Checks**

The most threatening data to the orthogonality of the evaluation cues and Information factor as independent variables concerned the significant effects of CI on three of the four measures of evaluation awareness. Since the three evaluation awareness measures which were affected showed significant or near significant effects for the Test cue, this
suggested the possibility that the presence or absence of contingency information somehow influenced the effectiveness of the Test cue.

However, the evidence for this possibility is less than overwhelming since the three measures of evaluation awareness showed incongruent relationships with the Information manipulation. That is, on the Dulaney item, No Information subjects indicated more effects of the Test, but on the Conformity and Gullibility attributions, it was the Information subjects who showed more effects of the Test. Since the best measure of evaluation awareness showed no effects for the Information, it was concluded that the most reasonable interpretation of the overall pattern of data was that the evaluation cues and Information manipulations produced relatively orthogonal effects on the predicted components of awareness.

Refined Predictions Regarding Evaluation Cues. The data from the impact of the evaluation cues on EA suggest that these cues function additively in producing evaluation awareness, since in no cases did any interactions approach significance. Because of this finding, a refinement of the H2, as suggested on page 79, is necessary. Rather than predicting an interaction between the cues on conditioning, it was predicted that the cues would be additive, as shown in Figure 4.
Type II: The Effects of Manipulations on Conditioning

For the analyses which follow in this section, subjects for whom no SD measure was available were dropped, yielding a total N of 322.

The first analysis performed involved a $2 \times 2 \times 2 \times 2 \times 2$ analysis with the three cue manipulations, high and low social desirability and sex of subject as the factors. This preliminary analysis was done because previous pilot work suggested the possibility of a sex of subject interaction with the Test cue. However, since the sex main effect was insignificant, as well as all the interactions of it with any of the other factors, it is not presented, and all the analyses were collapsed across sex.

The means for conditioning for the $2 \times 2 \times 2 \times 2$ analysis are presented in Table 7 and graphed in Figure 5, with significant differences indicated. The analysis of variance summary table is shown in Table 8. Since a four-factor interaction was obtained the data were analyzed according to simple main effects of the four factors (Winer, 1962), which are presented in Table 9. Evidence for each of the related predictions from these simple main effects are presented in order.

H1. It was predicted that the significant simple effects of contingency information in increasing conditioning would be most numerous when no evaluation cues were present, least
Fig. 4. Refined predicted interaction of independent variables on conditioning based on manipulation checks.
TABLE 7

OBTAINED MEANS FOR CONDITIONING WITH Ns INDICATED

<table>
<thead>
<tr>
<th></th>
<th>High SD</th>
<th></th>
<th>Low SD</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CI</td>
<td>NCI</td>
<td>CI</td>
<td>NCI</td>
</tr>
<tr>
<td>No Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Accomplice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>n=19</td>
<td>n=23</td>
<td>n=22</td>
<td>n=18</td>
</tr>
<tr>
<td></td>
<td>5.26</td>
<td>2.64</td>
<td>4.86</td>
<td>2.78</td>
</tr>
<tr>
<td>Test</td>
<td>n=17</td>
<td>n=18</td>
<td>n=24</td>
<td>n=19</td>
</tr>
<tr>
<td></td>
<td>3.58</td>
<td>3.44</td>
<td>5.08</td>
<td>1.94</td>
</tr>
<tr>
<td>No Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accomplice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>n=19</td>
<td>n=23</td>
<td>n=20</td>
<td>n=17</td>
</tr>
<tr>
<td></td>
<td>3.37</td>
<td>2.30</td>
<td>4.30</td>
<td>1.35</td>
</tr>
<tr>
<td>Test</td>
<td>n=20</td>
<td>n=21</td>
<td>n=21</td>
<td>n=20</td>
</tr>
<tr>
<td></td>
<td>1.65</td>
<td>.67</td>
<td>2.04</td>
<td>1.45</td>
</tr>
</tbody>
</table>
Fig. 5. Obtained means for conditioning with significant simple effects indicated.
<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contingency Information (CI)</td>
<td>228.139</td>
<td>1</td>
<td>228.139</td>
<td>36.195</td>
<td>.001</td>
</tr>
<tr>
<td>Test (T)</td>
<td>61.139</td>
<td>1</td>
<td>61.139</td>
<td>9.699</td>
<td>.001</td>
</tr>
<tr>
<td>Accomplice (A)</td>
<td>192.609</td>
<td>1</td>
<td>192.609</td>
<td>30.558</td>
<td>.001</td>
</tr>
<tr>
<td>Social Desirability (SD)</td>
<td>.960</td>
<td>1</td>
<td>.960</td>
<td>.152</td>
<td></td>
</tr>
<tr>
<td>CI X T</td>
<td>18.481</td>
<td>1</td>
<td>18.481</td>
<td>2.930</td>
<td>(.10)</td>
</tr>
<tr>
<td>CI X A</td>
<td>7.027</td>
<td>1</td>
<td>7.027</td>
<td>1.114</td>
<td></td>
</tr>
<tr>
<td>CI X SD</td>
<td>19.453</td>
<td>1</td>
<td>19.453</td>
<td>3.086</td>
<td>(.10)</td>
</tr>
<tr>
<td>T X A</td>
<td>19.851</td>
<td>1</td>
<td>19.851</td>
<td>3.149</td>
<td>(.10)</td>
</tr>
<tr>
<td>T X SD</td>
<td>2.643</td>
<td>1</td>
<td>2.643</td>
<td>.418</td>
<td></td>
</tr>
<tr>
<td>A X SD</td>
<td>2.501</td>
<td>1</td>
<td>2.501</td>
<td>.396</td>
<td></td>
</tr>
<tr>
<td>CI X T X A</td>
<td>1.339</td>
<td>1</td>
<td>1.339</td>
<td>.212</td>
<td></td>
</tr>
<tr>
<td>CI X T X SD</td>
<td>1.969</td>
<td>1</td>
<td>1.969</td>
<td>.312</td>
<td></td>
</tr>
<tr>
<td>CI X A X SD</td>
<td>1.191</td>
<td>1</td>
<td>1.191</td>
<td>.188</td>
<td></td>
</tr>
<tr>
<td>T X A X SD</td>
<td>1.050</td>
<td>1</td>
<td>1.050</td>
<td>.166</td>
<td></td>
</tr>
<tr>
<td>CI X T X A X SD</td>
<td>41.732</td>
<td>1</td>
<td>41.732</td>
<td>6.62</td>
<td>.05</td>
</tr>
<tr>
<td>Error</td>
<td>1,922.0</td>
<td>305</td>
<td>6.303</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 9

SIMPLE EFFECTS OF TEST X ACCOMPLICE X INFORMATION X SOCIAL DESIRABILITY INTERACTION ON CONDITIONING

I. Effect of Test

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Info-Accom-High SD</td>
<td>29.36</td>
<td>1</td>
<td>29.360</td>
<td>4.658</td>
<td>.05</td>
</tr>
<tr>
<td>Info-Accom-Low SD</td>
<td>50.69</td>
<td>1</td>
<td>50.692</td>
<td>8.042</td>
<td>.01</td>
</tr>
<tr>
<td>Info-No Accom-High SD</td>
<td>28.01</td>
<td>1</td>
<td>28.012</td>
<td>4.444</td>
<td>.05</td>
</tr>
<tr>
<td>Info-No Accom-Low SD</td>
<td>.45</td>
<td>1</td>
<td>.450</td>
<td>.071</td>
<td>NS</td>
</tr>
<tr>
<td>No Info-Accom-High SD</td>
<td>26.37</td>
<td>1</td>
<td>26.370</td>
<td>4.183</td>
<td>.05</td>
</tr>
<tr>
<td>No Info-Accom-Low SD</td>
<td>.099</td>
<td>1</td>
<td>.099</td>
<td>.0157</td>
<td>NS</td>
</tr>
<tr>
<td>No Info-Accom-High SD</td>
<td>6.19</td>
<td>1</td>
<td>6.194</td>
<td>.983</td>
<td>NS</td>
</tr>
<tr>
<td>No Info-Accom-Low SD</td>
<td>7.00</td>
<td>1</td>
<td>7.003</td>
<td>1.111</td>
<td>NS</td>
</tr>
<tr>
<td>Error</td>
<td>1922</td>
<td>305</td>
<td>6.303</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

II. Effect of Accomplice

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Info-Test-High SD</td>
<td>36.970</td>
<td>1</td>
<td>36.970</td>
<td>5.866</td>
<td>.05</td>
</tr>
<tr>
<td>Info-Test-Low SD</td>
<td>91.723</td>
<td>1</td>
<td>91.723</td>
<td>14.552</td>
<td>.001</td>
</tr>
<tr>
<td>Info-No Test-High SD</td>
<td>35.451</td>
<td>1</td>
<td>35.451</td>
<td>5.625</td>
<td>.05</td>
</tr>
<tr>
<td>Info-No Test-Low SD</td>
<td>3.265</td>
<td>1</td>
<td>3.265</td>
<td>.518</td>
<td>NS</td>
</tr>
<tr>
<td>No Info-Test-High SD</td>
<td>76.152</td>
<td>1</td>
<td>76.152</td>
<td>12.082</td>
<td>.001</td>
</tr>
<tr>
<td>Condition</td>
<td>Mean 1</td>
<td>SD 1</td>
<td>Mean 2</td>
<td>SD 2</td>
<td>t-value</td>
</tr>
<tr>
<td>----------------------------</td>
<td>--------</td>
<td>------</td>
<td>--------</td>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>No Info-Test-Low SD</td>
<td>2.383</td>
<td>1</td>
<td>2.383</td>
<td>.378</td>
<td>NS</td>
</tr>
<tr>
<td>No Info-No Test-High SD</td>
<td>1.216</td>
<td>1</td>
<td>1.216</td>
<td>.193</td>
<td>NS</td>
</tr>
<tr>
<td>No Info-No Test-Low SD</td>
<td>20.295</td>
<td>1</td>
<td>20.295</td>
<td>3.220</td>
<td>NS</td>
</tr>
<tr>
<td>Error</td>
<td>1922</td>
<td>305</td>
<td>6.303</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**III. Effect of Information**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean 1</th>
<th>SD 1</th>
<th>Mean 2</th>
<th>SD 2</th>
<th>t-value</th>
<th>df</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test-Accom-High SD</td>
<td>9.531</td>
<td>1</td>
<td>9.531</td>
<td>1.512</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test-Accom-Low SD</td>
<td>3.450</td>
<td>1</td>
<td>3.450</td>
<td>.547</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test-No Accom-High SD</td>
<td>.194</td>
<td>1</td>
<td>.194</td>
<td>.308</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test-No Accom-Low SD</td>
<td>97.856</td>
<td>1</td>
<td>97.856</td>
<td>15.525</td>
<td>.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Test-Accom-High SD</td>
<td>11.362</td>
<td>1</td>
<td>11.362</td>
<td>1.803</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Test-Accom-Low SD</td>
<td>86.371</td>
<td>1</td>
<td>86.371</td>
<td>13.703</td>
<td>.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Test-No Accom-High SD</td>
<td>67.611</td>
<td>1</td>
<td>67.611</td>
<td>10.727</td>
<td>.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Test-No Accom-Low SD</td>
<td>42.940</td>
<td>1</td>
<td>42.940</td>
<td>6.817</td>
<td>.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>1922</td>
<td>305</td>
<td>6.303</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**IV. Effect of Social Desirability**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean 1</th>
<th>SD 1</th>
<th>Mean 2</th>
<th>SD 2</th>
<th>t-value</th>
<th>df</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Info-Test-Accom</td>
<td>1.510</td>
<td>1</td>
<td>1.510</td>
<td>.240</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Info-Test-No Accom</td>
<td>22.331</td>
<td>1</td>
<td>22.331</td>
<td>3.543</td>
<td>.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Info-No Test-Accom</td>
<td>8.584</td>
<td>1</td>
<td>8.584</td>
<td>1.302</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Info-Test-No Accom</td>
<td>1.588</td>
<td>1</td>
<td>1.588</td>
<td>.252</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Info-Test-Accom</td>
<td>6.036</td>
<td>1</td>
<td>6.036</td>
<td>.958</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Info-Test-No Accom</td>
<td>22.331</td>
<td>1</td>
<td>22.331</td>
<td>3.543</td>
<td>.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Info-No Test-No Accom</td>
<td>.168</td>
<td>1</td>
<td>.168</td>
<td>.026</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Info-No Test-No Accom</td>
<td>8.957</td>
<td>1</td>
<td>8.957</td>
<td>1.421</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>1922</td>
<td>305</td>
<td>6.303</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
numerous when both evaluation cues were present, and of intermediate number when only one evaluation cue was present.

The evidence for this prediction was for the most part supportive. In both High and Low SD conditions, information was not effective in increasing conditioning when both cues were present, was sometimes effective when only one cue was present and was always effective in increasing conditioning when no cues were present. The contrary evidence arose from the instances in which the information was effective in the one cue condition. As can be seen in Figure 5, these occurred among Low SD rather than High SD subjects.

H2. The second prediction was that the significant simple effects of evaluation cues in decreasing conditioning would be more numerous in I conditions than in NI conditions.

The evidence for this prediction is listed separately for each evaluation cue. First of all, with respect to the test cue, Figure 5 shows that it was effective in reducing conditioning in the NI conditions only once (that was in the Accomplice-High SD condition), but in the I conditions, the test was effective three of the four times (everywhere except the No Accomplice-Low SD condition). This series of differences supports the prediction, and the instances of when it was ineffective is consistent with the hypothesis. That is, the test made no difference in the I condition only when the accomplice was absent, among Low SD subjects, but it was effective in reducing conditioning in the NI condition only when the Accomplice was present, among High SD subjects.
The evidence for the Accomplice simple effects were also supportive, where it can be seen that the accomplice was effective in reducing conditioning three of the four times when information was given (everywhere except in the Low SD-No Test condition, which is where the least effect was expected). Similarly, it was effective in reducing conditioning only once in the NI condition, and that was in the High SD-Test condition, where the strongest effects were expected.

In general the evidence for the predicted additivity of the evaluation cues was not supported in that, as Figure 5 indicates, the cues interacted with respect to their effects on conditioning. Among HSD subjects in the NI condition, the Accomplice was effective only when the Test was present, and the Test was effective only when the Accomplice was present. Among Low SD subjects in the I condition, this pattern also occurred.

H3. The third prediction was that the significant simple effects of the social-approval factor would be most numerous when information and evaluation cues were present, least numerous when information and evaluation cues were absent, and of intermediate number when some but not all of the cues were present.

In general the data did not support this specific hypothesis. As Table 9 indicates, the SD factor did not produce any significant simple effects, and approached significance ($p < .10$) in only two cases. The first was in the
Information-Test-No Accomplice condition in which High SD subjects did show less conditioning than Low SD subjects, as predicted. However, the other case was incongruent with the prediction, where, in the No Information-Test-No Accomplice condition, High SD subjects showed more conditioning than Low SD subjects. Neither of these conditions were those in which the largest effects of SD were expected (when all cues present) and since one was in the wrong direction, the evidence is not at all supportive.

However, an alternative way of viewing the simple effects data does suggest support for the notion that the effects of the cues differ for Low SD and High SD subjects. That is, in line with the conceptualized functioning of SD, one would expect the evaluation cues to be more effective among High SD than among Low SD subjects, and Figure 5 shows that this was clearly the case. The accomplice was effective in reducing conditioning among High SD subjects everywhere except in the No Test-No Information condition, whereas it was effective among Low SD subjects only in the Test-Information Condition. Similarly, the Test was effective in reducing conditioning among High SD subjects everywhere except in the No Accomplice-No Information condition, whereas among Low SD subjects, it was only effective in the Accomplice-Information condition. This pattern fits very well with the notion that SD interacts with the effects of the evaluation cues.
However, with regard to the information cue, the simple effects do not seem to fit this pattern. In fact, just the opposite effect occurred. That is, the information was effective in enhancing conditioning in Low SD subjects three of four times, everywhere except when both evaluation cues were present, whereas it was effective only once among the High SD subjects (that was when neither evaluation cue was present). These results are considered further in the discussion section.

Proportion of Variance Accounted For

As presented in Table 8 (p. 98), where the design is analyzed into the sources of variance, the total proportion of variance accounted for with the manipulations and social desirability measure is .2379. This estimate was derived by adding all the significant mean squares in the analysis and dividing by the total mean square.

Type III: The Study of Awareness

The intercorrelations between the awareness measures collected from the post-experimental questionnaire are presented in Table 10. As will be shown, the data supported the prediction that hypothesis awareness is necessary for the operation of evaluation cues, but the multiplicative network derived from Dulaney's theorizing received little empirical support.
TABLE 10

INTERCORRELATIONS AMONG AWARENESS MEASURES AT SELECTED LEVELS OF HYPOTHESIS AWARENESS

<table>
<thead>
<tr>
<th></th>
<th>CA</th>
<th>HA</th>
<th>EA</th>
<th>MC-H</th>
<th>MC-E</th>
<th>HA X MCH</th>
<th>HA X MCE</th>
<th>DV</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>.709</td>
<td>.004</td>
<td>-.141</td>
<td>-.264</td>
<td>.430</td>
<td>.579</td>
<td>.471</td>
<td></td>
</tr>
<tr>
<td>HA</td>
<td>.044</td>
<td>-.192</td>
<td>-.271</td>
<td>.699</td>
<td>.856</td>
<td>.479</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA</td>
<td>-.077</td>
<td>-.102</td>
<td>-.019</td>
<td>-.007</td>
<td>.332</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCH</td>
<td></td>
<td>-.227</td>
<td>.151</td>
<td>.308</td>
<td>-.247</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCE</td>
<td></td>
<td></td>
<td>.411</td>
<td>-.033</td>
<td>-.308</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HA X MCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.724</td>
<td>.195</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HA X MCE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.314</td>
<td></td>
</tr>
<tr>
<td>DV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ df = 100 \]

Critical values: \(.10=.164; .05=.195; .01=.254 \)

Code Key:

<table>
<thead>
<tr>
<th>Item</th>
<th>Construct</th>
<th>Items from post-experimental questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>Contingency Awareness</td>
<td>#6 and #7</td>
</tr>
<tr>
<td>HA</td>
<td>Hypothesis Awareness</td>
<td>#8 and #9</td>
</tr>
<tr>
<td>EA</td>
<td>Evaluation Hypothesis</td>
<td>#11 and #12</td>
</tr>
<tr>
<td>MC-H</td>
<td>Motivation to Comply (Help)</td>
<td>#15</td>
</tr>
<tr>
<td>MC-E</td>
<td>Motivation to Comply (Expect)</td>
<td>#14</td>
</tr>
<tr>
<td>HA X MCH</td>
<td>Hypothesis Awareness multiplied by Motivation to Comply--Help</td>
<td></td>
</tr>
<tr>
<td>HA X MCE</td>
<td>Hypothesis Awareness multiplied by Motivation to Comply--Expect</td>
<td></td>
</tr>
<tr>
<td>DV</td>
<td>Dependent Variable (Conditioning)</td>
<td></td>
</tr>
</tbody>
</table>
Evidence for the Central Role of Hypothesis Awareness

As can be seen from Table 11, which presents the correlations between the evaluation cue manipulations, evaluation awareness, and the motivation to comply measures with conditioning at selected levels of hypothesis awareness, hypothesis awareness was critical in mediating the effect of evaluation cues. When hypothesis awareness was essentially absent (HA less than 3), the correlations between conditioning and the evaluation cues and awareness were not significantly different from zero. However, when hypothesis awareness was present (HA more than 5), the correlations were significant. This pattern of findings suggests that the effects of evaluation apprehension depend upon the prior presence of hypothesis awareness.

The simple effects of the evaluation cues at I vs. NI conditions would lead one to expect this kind of pattern. However, the interaction from the analysis of variance data alone would not enable the specification of which type of awareness plays the more central role, since it could also be inferred that the effects of contingency information depend upon the prior reception of evaluation cues. The correlational data derived from examining the effects of contingency awareness do enable the delineation of an asymmetrical relationship, however, since the converse analysis did not produce the same pattern of effects. That is, hypothesis awareness played a significant role at selected levels of evalu-
### TABLE 11

**INTERCORRELATIONS AMONG AWARENESS MEASURES**

#### I. When Hypothesis Awareness = 6 or 7 (n = 252)*

<table>
<thead>
<tr>
<th></th>
<th>Test</th>
<th>Accom</th>
<th>EA-MC</th>
<th>EA-Essay</th>
<th>EA-Gull</th>
<th>EA-Cont</th>
<th>MCE</th>
<th>MCH</th>
<th>SD</th>
<th>DV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test</strong></td>
<td>-0.071</td>
<td>0.171</td>
<td>0.297</td>
<td>0.164</td>
<td>0.000</td>
<td>0.156</td>
<td>0.027</td>
<td>0.038</td>
<td>0.211</td>
<td></td>
</tr>
<tr>
<td><strong>Accom</strong></td>
<td>0.033</td>
<td>0.103</td>
<td>0.015</td>
<td>0.080</td>
<td>0.050</td>
<td>0.089</td>
<td>0.069</td>
<td>0.283</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EA-MC</strong></td>
<td>0.132</td>
<td>0.288</td>
<td>0.292</td>
<td>0.139</td>
<td>0.018</td>
<td>0.078</td>
<td>0.131</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EA-Essay</strong></td>
<td>0.311</td>
<td>0.294</td>
<td>0.106</td>
<td>0.114</td>
<td>0.060</td>
<td>0.448</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EA-Gull</strong></td>
<td>0.372</td>
<td>0.138</td>
<td>0.060</td>
<td>0.034</td>
<td>0.226</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EA-Cont</strong></td>
<td>0.016</td>
<td>0.094</td>
<td>0.040</td>
<td>0.181</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MCE</strong></td>
<td>0.431</td>
<td>0.082</td>
<td>0.217</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MCH</strong></td>
<td>0.054</td>
<td>0.235</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.041</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DV</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### II. When Hypothesis Awareness = 0, 1 or 2 (n = 71)**

<table>
<thead>
<tr>
<th></th>
<th>Test</th>
<th>Accom</th>
<th>EA-MC</th>
<th>EA-Essay</th>
<th>EA-Gull</th>
<th>EA-Cont</th>
<th>MCE</th>
<th>MCH</th>
<th>SD</th>
<th>DV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test</strong></td>
<td>0.095</td>
<td>0.042</td>
<td>0.211</td>
<td>0.018</td>
<td>0.015</td>
<td>0.023</td>
<td>0.000</td>
<td>0.011</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td><strong>Accom</strong></td>
<td>0.000</td>
<td>0.148</td>
<td>0.106</td>
<td>0.161</td>
<td>0.141</td>
<td>0.099</td>
<td>0.242</td>
<td>0.153</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EA-MC</strong></td>
<td>0.097</td>
<td>0.084</td>
<td>0.096</td>
<td>0.109</td>
<td>0.365</td>
<td>0.099</td>
<td>0.059</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EA-Essay</strong></td>
<td>0.249</td>
<td>0.274</td>
<td>0.229</td>
<td>0.061</td>
<td>0.022</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EA-Gull</strong></td>
<td>0.561</td>
<td>0.059</td>
<td>0.100</td>
<td>0.105</td>
<td>0.039</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EA-Cont</strong></td>
<td>0.041</td>
<td>0.029</td>
<td>0.010</td>
<td>0.053</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MCE</strong></td>
<td>0.351</td>
<td>0.086</td>
<td>0.168</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MCH</strong></td>
<td>0.026</td>
<td>0.080</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.096</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DV</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Critical values: df = 100; .10=.164; .05=.195; .01=.254

**Critical values: df = 70; .10=.195; .05=.232; .01=.303
tion awareness. When subjects were aware of a negative evaluation (EA at -2 or -1 on the essay measure), the correlation between hypothesis awareness and conditioning was .3208. When evaluation awareness was in the positive direction (EA at +1 or +2 on the essay measure), the correlation was .5460, and when subjects indicated no awareness (EA at 0), the correlation was .4817. This pattern is exactly as one would expect, in that the positive awareness should enhance the tendency for subjects to cooperate with their hypothesis awareness, and the negative direction should mitigate this tendency. Overall, these data form a remarkably clear pattern in suggesting that the effects of evaluation awareness depend upon the subject knowing what he is supposed to do in the experiment, and that the direction of evaluation awareness enhances or detracts from the subjects cooperation with that knowledge.

Evidence for the Dulaney Model

On the other hand, the more elaborate theoretical network derived from the Dulaney paradigm did not receive much support. The multiplied HA x MC term did not bear a greater relationship to the conditioning response than did the awareness of contingencies or evaluation. As can be seen from Table 11, the HA x MC terms bore little relationship to the dependent variable (.195 for HA x MC_H and .314 for HA x MC_E). These correlations were both lower than the relationship of CA to the dependent variable (.471) as well
### TABLE 12

**CORRELATIONS BETWEEN SD AND EA AT SEPARATE EVALUATION CUE CONDITIONS**

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>NI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No Test</strong></td>
<td>-.0207</td>
<td>.0982</td>
</tr>
<tr>
<td><strong>Test</strong></td>
<td>-.2941</td>
<td>.1426</td>
</tr>
<tr>
<td><strong>No Accom</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Accom</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test</strong></td>
<td>-.2031</td>
<td>.1294</td>
</tr>
<tr>
<td></td>
<td>-.1495</td>
<td>.1279</td>
</tr>
</tbody>
</table>

**df = 38**

**Critical values**

- .10 = .275
- .05 = .325
- .01 = .381
I. When HA X MCE out:

```
CA --- HA X MCE --- DV
       .437

EA     .005     .343
```

II. When HA X MCH out:

```
CA --- HA X MCH --- DV
       .373

EA     .002     .354
```

Fig. 6. Correlations between model components when HA X MC partialled out.
as the relationship of EA to the dependent variable (0.332).

Other predictions based on this model also failed to be supported. The partialling of the multiplicative HA x MC terms did not reduce the relationship between the DV and EA or the DV and CA to zero. As Figure 6 (p. 109) shows, partialling left significant relationships between conditioning and contingency awareness (0.437 for MC_E and 0.373 for MC_H) as well as between conditioning and evaluation awareness (0.343 for MC_E and 0.354 for MC_H).

Thus it appears that the effects of contingency and evaluation awareness do not operate through a mediating process involving a multiplicative relationship of hypothesis awareness and motivation to comply. Reasons for the failure of this model to adequately describe the data, as well as the discrepancy between the model and the data for the central role of hypothesis awareness are dealt with in the discussion.

Proportion of Variance Accounted for by the Correlational Data

When the correlations between the awareness measures and the conditioning scores were subjected to a multiple regression analysis, the result was a multiple R^2 of 0.402, indicating that all of the awareness measures together accounted for about 40 percent of the variance.

Since experimental manipulations accounted for about 24 percent of the variance, a step-wise multiple regression analysis of all of the awareness measures and manipulations
was run to estimate the total proportion of unique variance accounted for in the conditioning scores by all of the variables in the study. The multiple $R^2$ yielded from this analysis was .449.
VI. DISCUSSION

The pattern of findings just described indicates that many, though not all, of the predictions were supported. The manipulation checks suggested that the hypothesis and evaluation cues were successfully manipulated. The major predictions concerning the interaction of cues with subject need for approval were mostly supported, with the exception of a few unanticipated series of effects which necessitate some explication below. Finally, the study of post-experimental awareness suggested that hypothesis awareness is a critical component in mediating the effects of evaluation cues, but that the evidence for the Dulaney model was not supportive.

Several issues emerge from the pattern of results obtained in this investigation. The question of the relative importance of subject roles, difficulties in measuring evaluation awareness and its relationship to social desirability, problems in interpreting the motivation to comply measures, the controversy regarding the degree to which subjects engage in conscious decision making in this paradigm, and recommendations for future research are discussed.

Experimental Evidence for Subject Roles

The pattern of data described for the convergent and discriminant validity checks of the major independent variables of evaluation cues and contingency information suggests
that in this paradigm, the two types of cues were manipulated successfully. This allowed for the empirical test of the relative importance of the good and apprehensive roles, which were found to be a systematic function of the number and type of cues found in the experimental situation. That is, contingency information clearly increased conditioning, indicating support for the good-subject role, except when evaluation cues were numerous and subjects were high in need for social approval. On the other hand, the evaluation cues decreased conditioning, except when no information was given and subjects were low in need for social approval. On the whole, the effects of the cues interacted in a very consistent pattern with the social desirability factor to suggest that cooperation or resistance with the demands of an experiment depend upon the cues in the situation as well as on individual characteristics of the subjects. In this sense, the term "role" seems an inappropriate label for the extent to which subjects cooperate, since it suggests that it might be trans-situational as well as independent of the unique characteristics of the subject. Instead, this research has shown that cooperation depends very clearly on a complex interaction between specific features of the experimental situation and aspects of the individual in that situation.

Overall, the predictions were well supported. However, there were a number of inconsistencies in the data which deserve closer attention.
Relationship of SD Factor to Manipulated Cues

One major series of disconfirming findings within the four factor interaction occurred with the effects of SD. In contradiction to the hypothesis that the SD factor would be more effective as more cues were introduced, what in fact occurred was that the SD factor showed no significant effects at all. Those places where it approached significance were not in the conditions where the strongest effects were predicted, nor were they always in the predicted direction.

However, as reported in the results section, consideration of the simple effects of the evaluation cues at different SD levels shows that a pattern of results was obtained that was clearly consistent with the underlying conceptualization of the role of SD in the present paradigm. That is, evaluation cues were more effective for High SD subjects than for Low SD subjects, and the instances in which they were ineffective in each case were those which would be predicted when this perspective is taken.

The fact that the Information simple effects showed an opposite pattern to the evaluation cue effects seem contradictory to the above pattern. That is, the simple effects of information showed that Low SD subjects were more often affected by the information than were High SD subjects. This finding seems to also contradict past research which suggests that High SD subjects are more cooperative in a conditioning experiment (Crowne and Strickland, 1962).
However, a look at the conditions where High and Low SD subjects showed significantly more conditioning when contingency information was present suggests a plausible explanation. Specifically, High SD subjects were only affected by the Information when both the Test and Accomplice were absent, and Low SD subjects were affected by the Information everywhere except when both the Test and Accomplice were present. This pattern is very similar to the effects of Information at different evaluation cue levels, since the Information was most effective when the cues were absent, and least effective when they were both present. It seems very plausible, then, that the SD factor could be operating as a "chronic" evaluation cue, that is, a heightened propensity to become evaluation aware with which the subject enters the laboratory. If this were the case, rather than facilitating the effects of the Information cue, it would be expected that the SD factor would mitigate against it, just as the other evaluation cues tended to do. This overall pattern suggests a very consistent picture, supporting the role of SD in interaction with the cue manipulations.

Thus, the series of findings from the simple effects of the cues at different SD levels seems to contradict those obtained for the simple effects of SD at different cue levels. One possible explanation for this incongruency involves the relative weakness of the SD factor in comparison to the strength of the cue factors. That is, the various cues produced very powerful effects on behavior, as shown by the main
effects from Table 8 (p. 98), but the 5D factor did not. One reason for the relative impotency of the 5D factor could involve the operationalization of SD which very likely hindered the possibility of it showing significant effects.\(^3\) Inspection of Figure 2 shows that the distribution of SD scores was fairly leptokurtic. Over half of the subjects scored within four points of each other, and were thus divided into High and Low SD conditions on the basis of a fairly arbitrary criterion (median split). If the definition of the SD factor had involved trichotomization rather than dichotomization, 140 scorers who ranged within 3 points would have been eliminated from the High and Low conditions and a more powerful SD factor would have been likely to result. This procedure would most likely increase power by increasing between cell variance relative to within cell variance, since the loss of degrees of freedom with large N research would not change the error term a great deal. This possibility could be easily tested by future research, and if the middle third was retained as a middle level, the possibility for checking curvilinear functions of the SD factor would also be gained. This trichotomization was not undertaken in the present research since a three level factor was felt to add too much complexity to an already complex design. However, the present findings suggest this procedure would have substantial payoff for future research.

\(^3\)One difficulty with this explanation involves the fact that the SD factor did not approach significance in the multiple cue conditions where the strongest simple effects were predicted to occur.
This interpretation suggests that the SD scale is validly differentiating subjects based on their tendency to pick up and respond to evaluation cues, and rests on the assumption that the High SD subjects are more evaluation aware than the Low SD subjects. Interestingly, this link in the rationale does not receive support from the awareness data, since the correlation between the SD measure and evaluation awareness as measured by the essay item was -.0076, which is not significant. (The SD measure also failed to correlate significantly with the other three measures of evaluation awareness, all of which were deemed poorer measures on the basis of convergent and discriminant validity data.)

Since the essay item received empirical support for its convergent and discriminant validity, the lack of relationship between it and the SD measure is a troublesome incongruency. Perhaps the SD scale is at fault. What empirical evidence exists for the SD scale as a valid measure of the tendency to be affected by evaluation cues?

Fortunately, this measure of need for approval has been subjected to considerable construct validation testing. In the 1964 text describing the development and empirical investigations of the scale, Marlow and Crowne report that the SD scale was predictive of behavior in a wide variety of social approval situations. For example, compared to subjects scoring low in need for social approval, high SD subjects were found to 1) report more favorable attitudes toward
a boring experimental task to an inquiring experimenter, 2) conform more to inaccurate perceptions of auditory stimuli given by a group of accomplices in an Asch-type conformity situation, 3) give more common word associations, 4) report less taboo words in a perceptual defense paradigm, 5) show more attitude change after role-playing, 6) be more defensive against hostility by being more influenced by a euphoric accomplice after instigation to aggression in a Schachter-Singer type of waiting experiment, 7) terminate psychotherapy earlier even though their therapists regard them as more defensive and afraid of social rejection, and 8) show higher need for affiliation on a projective measure, but be less liked by peers. More recent research by others has replicated and extended these findings to other social approval situations (Hollender, 1969; McLaughlin & Hewitt, 1972; Meisels & Ford, 1965; Miller, Doob, Butler, & Marlow, 1965; Posavaac, 1971; Salman, 1962).

Most relevant to the present research is the chapter which relates the findings from verbal conditioning studies where it is reported that high need for approval subjects showed greater verbal conditioning in the Greenspoon paradigm, but only when the reinforcement involved social approval. This finding has been replicated by others (Epstein, 1964). Thus, evidence from a wide variety of research paradigms converges with the conditioning experiment to suggest that those who score high on the SD scale are more concerned about socially approved behavior in an experiment, and monitor their
behavior accordingly. This network of findings would lead one to expect this scale measures individual differences in the propensity for evaluation apprehension in the sense in which Rosenberg originally defined the term.

The studies in which Rosenberg has used the scale are less supportive, however. While the data he reports with his picture rating paradigm (1969) are somewhat sketchy for the SD measure, it appears that in one study, evaluation cues did seem to be more effective for High SD subjects, at least when the cues suggested that mature people tend to like the faces. When the cue involved disliking, however, Low SD subjects did not confirm the relationship. However, the most serious problems involved in the results from a second study, where subjects were told that bored and inefficient performance on a very simple experimental task was a correlate of maturity. In this case, Low SD subjects produced lower scores than High SD subjects, who performed similarly to a control group without the evaluation cue. While Rosenberg interpreted this as due to credibility problems of the manipulation, he provided no such empirical evidence, and we are forced to conclude that the SD scale simply did not function as one would have predicted on the basis of the constructs involved.

One reason for this might relate to the present research in which the SD factor also functioned less than perfectly. That is, in contrast to most successful demonstrations of differences between High and Low SD subjects to
approval cues, both Rosenberg's and the present paradigm gathered data from subjects in a group situation. One would expect conformity to socially approved norms to be greatest for the individual acting alone in the experiment, since the presence of other subjects could weaken the effects. For example, to the extent that High SD subjects could use other subjects as reference information, or to allay evaluation anxiety, one would expect the effects of the SD factor to be weakened. Future research could test this possibility by running the same paradigm with individual subjects and comparing the effects of the SD factor when subjects are run in groups.

In sum, it seems that there are some inconsistencies in findings with the SD scale, most notably from the laboratory of Rosenberg himself, but that overall a large body of evidence can be marshalled for its support. Given the systematic way in which the cues interacted with the scale in the present study, it seems difficult to dismiss it as an inadequate measure.

On the other hand, this research went considerably further than others in developing a measure of evaluation awareness which satisfied both interjudge reliability and convergent and discriminance validity criteria. Since the measure also correlated with the dependent variable in the predicted direction, it appeared to be an adequate measure of evaluation awareness. Thus, evidence for the validity of both the EA and SD measure is convincing, and the lack of
relationship between them in the present study is puzzling.

One possibility for this finding is the manner in which the correlation was calculated, that is, across all subjects in all conditions. One would expect that the SD factor might bear different relationships to EA in different cue conditions. Negative correlations would be expected in conditions where the E cue was strongest. In other conditions one would expect positive and zero correlations, and collapsing across all conditions could thus produce no correlation even though meaningful relationships existed. To check this possibility, the correlation between SD and EA was recalculated at separate cue conditions, and these coefficients are presented in Table 12. A pattern of positive and negative correlations did appear, although not the pattern which would be predicted from the above rationale. Negative correlations were found throughout the Information conditions, whereas positive correlations appeared when information was absent.

This finding suggests that the role of SD on mediating the effects of evaluation cues may depend upon the relative saliency of hypothesis cues. When the experimenter's hypothesis is clear, High SD subjects are more likely to discern the negative implication of conforming to it, but when the hypothesis is not clear, conditioning is not more likely to be discerned as an instance of conformity by High SD subjects.

This is an interesting pattern of data since it suggests that the effects of cues on awareness may take different functions depending upon the presence of other factors. In
the current study, the investigation of the cues on awareness was undertaken to provide manipulation checks of both convergent and discriminant validity and as such went a great deal further than previous research establishing validity of the manipulations. However, these analyses could be limited in that they were basically main effect analyses, that is, based on single factor designs. What the above finding suggests is that it is possible that interactions might emerge if evaluation cues and information factors as well as the SD factor were employed in the same design on the measures of awareness. If such interactions did emerge, the information about the effects of these cues on awareness would be considerably richer.

This methodology could help to answer a beguiling question which has repeatedly caught the attention of researchers in this field, and that is whether these sorts of evaluation cues affect subject behavior by affecting the subject's perception of the situation, or merely his willingness to cooperate with his perceptions. Since it has already been shown that the evaluation cues did not influence the subject's hypothesis awareness or contingency awareness, part of this question has already been tentatively answered in this research, at least with respect to main effect analyses. However, since this correlational analysis shows that information does affect the way in which the subject's need for approval relates his evaluation awareness, it could be that a more elaborate analysis of the awareness measures might indicate ways in which
the subject's need for approval mediates perception. Thus, there seems to be the possible distinction between the chronic definition of evaluation cues in terms of SD, which does interact with information cues to affect evaluation awareness, and the acute definition of evaluation cues in terms of the test and confederate which does not interact with information cues to affect evaluation awareness. This possibility could be checked by employing the four factor analysis on the various measures of awareness from the post-experimental questionnaire and comparing these results to the four factor analysis on awareness.

Evidence for the Multiplicative Model of Awareness

The evidence from the experimental data suggested that the hypothesis and evaluation cues did operate in an interactive manner in that hypothesis awareness enhanced the effects of evaluation cues, and that when evaluation awareness was especially strong, it cancelled the effects of hypothesis cues. The evidence from correlational data gathered from selected samples of subjects at high and low levels of hypothesis awareness further indicated that this interaction was produced by an asymmetrical relationship between hypothesis and evaluation awareness, hypothesis awareness playing a more central role by being necessary for the effectiveness of evaluation cues, while the converse was not true. That is, at selected levels of evaluation awareness, and especially when evaluation awareness was zero, hypothesis awareness still
enhanced conditioning, whereas when hypothesis awareness was at zero, evaluation awareness did not relate to conditioning. Thus, the correlational and experimental data converge to verify a model of subject cooperation in which the subject becomes aware of the experimenter's hypothesis and then decides to cooperate with it on the basis of evaluation cues present in the situation.

Unfortunately the multiplicative model of awareness derived from Dulaney's Theory of Propositional Control failed to complement the above set of findings. The HA x MC term did not correlate more highly with the dependent variable than components of the term, or more highly than did the exogenous components, CA and EA. What explanations exist for this discrepant set of findings?

The Dimension of Hypothesis Awareness

Since the subsample analysis involved the dichotomization of hypothesis awareness, whereas the partialling analysis involved the use of the entire dimension, it is possible that the use of the entire dimension introduced error into the system, error which was unnecessary and clouded the true relationships from view. A controversy has previously arisen over the proper dimensionalization of hypothesis awareness, and Page (1969) has dichotomized the component, basically because of the bimodal distribution which the term has taken. As can be seen from Table 13, the distribution which both contingency and hypothesis awareness took in the present study was far from normal although not necessarily very bimodal.
TABLE 13

DISTRIBUTIONS OF CONTINGENCY AND HYPOTHESIS AWARENESS

<table>
<thead>
<tr>
<th></th>
<th>N for CA</th>
<th>N for HA</th>
</tr>
</thead>
<tbody>
<tr>
<td>-6</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>-5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-4</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>-3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>-2</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>-1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>0</td>
<td>50</td>
<td>26</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>31</td>
</tr>
<tr>
<td>5</td>
<td>39</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>243</td>
<td>252</td>
</tr>
</tbody>
</table>
In that both components were highly skewed to the left (with 63 percent and 65 percent of the subjects scoring full awareness on CA and HA, respectively), the assumption of normal distribution for a Pearson correlation was not met, and simplification of the distribution by dichotomization seemed appropriate.

Further, the original rationale for the multiplied HA x MC component rested largely on the conceptualization of these terms when they were scored as zero. That is, if either term were zero, the resulting conditioning should also be zero. Since the dimensionalized HA component rarely reached a zero level, even though scores of 1 or 2 would indicate little or no awareness, HA was dichotomized into scores of 0 and 1.

Hence, the HA x MC terms were recalculated using a dichotomized HA term where subjects scoring below 3 were assigned a score of 0 and subjects scoring above 5 a score of 1. These new terms (HA^ x MC_H and HA^ x MD_E) produced a pattern of partials which are indicated in Figure 7. As can be seen, this rescoring did little to improve the empirical status of the model. The new terms only correlated .2793 and .3735 with the dependent variable, and partialling them from the CA-DV relationship still left it at .3579 and .4123 (for MC_H and MC_E, respectively).

Thus, while the dichotomization of HA improved somewhat the role of the multiplicative term in the model, the size of these effects were not enough to offer support for the causal
Fig. 7. Correlations between model components when HA scored as a dichotomy and resulting HA X MC term partialled out.
model. Apparently, the failure of the model to attain empirical support was not due to the scoring of hypothesis awareness.

The Validity of the Motivation to Comply Measures

Since the problem did not seem to be located in the scoring of hypothesis awareness, the next suspect was the motivation to comply measure. Indeed, previous research in other laboratories suggests that motivation to comply may be an extremely difficult construct to measure.

The first example of these problems appears in a thesis by one of Dulaney's students (Bottom, 1972) where the attempt to manipulate MC resulted in a nonconfirmation of Dulaney's theory. The study used a conditioning task in which color choices were the responses to be conditioned and the manipulation of MC was attempted by informing each subject that "so far we have learned that in this task the more normal, the less neurotic the subjects are, the more they very much want to choose whatever color they think they are supposed to choose [p. 45]." While Dulaney described this attempt to manipulate MC as one that "failed rather badly" (Dulaney, personal communication, 1973), it seemed from the analysis of the post-experimental reports that this was not precisely the case. Subjects were asked "When choosing one kind of color or another over the last 20 trials, how much did you want to choose whatever you might have thought you were supposed to choose?" and responses were scored from +1 (want
very much) to -1 (very much did not). The data for this item show that the manipulation was indeed successful in that the people told that "it is neurotic not to want to do what's expected" had a mean score of .61; those told the opposite, that "it is neurotic to want to do what's expected" had a mean score of .07; and those told that less neurotic people don't care" (a neutral condition) had a mean of .43. This difference was significant beyond the .001 level, indicating that this information did create different reports of motivation to comply with the experimental hypothesis.

However, even though the manipulation was successful, it was described as unsuccessful because using the reports of MC in the formula

\[
\left( \frac{\text{Contingency}}{\text{Awareness}} \right) \times \left( \frac{\text{Subjective Value of Reinforcement}}{\text{Behavioral Hypothesis}} \right) \times \left( \frac{\text{Motivation to Comply}}{\text{Motivation}} \right)
\]

did not increase the accuracy of predicting either behavioral intention or the dependent variable. In fact, the prediction system was more successful without the MC term than it was when it was included in the formula. Because of this, it was concluded that the true value of MC might have been 1, suggesting that MC might not have really been manipulated. However, this is a very weak explanation, since subjects did
report different MC values across conditions which differed in the expected direction. The conclusions seem unsatisfactory, but bear some relationship to a further study of MC by Uleman (1971b).

In this experiment, the attempt was made to heighten MC by telling subjects that good performance in a verbal conditioning task would be rewarded with an important position in a future group decision hierarchy. Furthermore, chronic levels of MC were measured with a Thematic Apperception type of need for Influence and Need for Power measure (Uleman, 1966). The results of this study show that the motivational variables affected conditioning, but not through the operation of Behavioral Intention, as Dulaney had specified. Instead, subjects who reported they wanted to get the reinforcement conditioned faster than those with low MC, even when the behavioral intention was zero. That is, the relationship between MC and DV was positive without the mediating influence of behavioral intention, suggesting that motivation to comply may not be so obviously anchored in awareness as are other aspects of the verbal conditioning situation such as contingency awareness and behavioral hypothesis.

Uleman's explanation for this unexpected finding involved the "self-perception phenomenon" which Bem (1967) has offered to explain self-knowledge about internal processes. Briefly, the position states that self-attributions about such things as attitudes and motives are based on our own external behavior more than internal processes which we can directly
experience. Thus, when subjects are asked about contingency awareness or hypothesis awareness, they can refer to external events in the environment and the awareness of these events is easy to report. However, when they are asked about purely internal events, such as intentions and motivations, they have no direct external referent, and must use a variety of potentially artifactual information to come up with an inference for the experimenter. In this particular case, Uleman suggests that the positive MC-DV relationship when behavioral intention was zero may have been caused by subjects with moderate MC who conditioned, but inferred little behavioral intention because they did not fully condition.

While in the present study it was not possible to check the operation of MC through behavioral intention, there are other reasons to suspect that it did not measure motivation in a valid manner. In the present study, the situational manipulation of evaluation cues should have affected motivation to comply through the mediating effects on EA. However, the relationship between these variables was quite low, as shown in Table 10 (p. 104). Evaluation awareness correlated only −.0779 with MC_H and −.1021 with MC_E. These correlations indicate that whatever the MC measures were doing, they were not measuring the motivation to cooperate which was hypothesized to be mediated through evaluation awareness.

While the lack of correlation between evaluation awareness and the MC terms is a threatening finding to the validity of the MC measures, the most devastating evidence
regards the low correlation between the MC measures and the dependent variable among hypothesis aware subjects. As indicated in Table 11 (p. 106), these correlations were only .2171 for $MC_E$ and .2349 for $MC_H$. While these correlations are significant, the fact that they are not higher suggests that subjects were not indicating their cooperation, even when they had full knowledge of the demands, and were conditioning. Even subjects who conditioned fully (scores on the dependent variable at 6 or 7) only indicated a mean cooperation of 2.95, and the standard deviation among this group of 1.627 on the $MC_E$ measure shows that they are considerably variable in their reports of cooperation. (The mean and standard deviation for the $MC_H$ measure were similar, 2.87 and 1.743, respectively.)

These relationships between MC and EA do contradict those reported by Page (1971) in a study of evaluation apprehension on verbal conditioning. He gave a personality test dealing with conformity as the evaluation cue manipulation, and measured evaluation awareness by asking subjects "what was the purpose of the personality test you took at the beginning?" While his measure of MC was not clearly spelled out, he reported positive effects. While only one out of his 14 EA subjects reported cooperation, 5 of his 8 non-evaluation aware subjects did, producing a significant relationship between EA and MC in his study. Further, the relationship between MC and the presence or absence of the personality test approached significance ($p < .10$) where twice as many subjects
in the no test group reported cooperation than did those in the test group.

The reason for the discrepancy between Page's study and the present one regarding the validity of the MC term remains to be tested. An obvious possibility is the difference in experimental paradigms, the verbal conditioning allowing a subject more responses, and thus more data on which to infer his cooperation with the hypothesis. It could be that the one response aspect of the classical conditioning task does not allow the subject enough information about his own behavior to extract an estimation of his degree of compliance. This may be the reason why the MC term has received partial support from Dulaney (1968) although in this case, the attempt was not made to manipulate motivation.

It could also be that the precise wording of the MC measures may significantly affect the responses, although the two variations used in the current research produced similar sets of data throughout many different analyses. Without knowing the exact wording of Page's item, it is difficult to know what subjects were reporting about themselves. However, since the $MC_F$ measure of the present study was precisely the one used by Dulaney, it is difficult to attribute all of the failure to this problem.

In any case, it seems reasonable that the MC measures would be among the most difficult components of post-experimental awareness to interpret, and even Page, who has been relatively successful in dealing with this term has
repeatedly acknowledged his hesitations in making inferences from it (1970, 1971, 1973). Page notes that many subjects in his studies indicated their reservations about answering the item, and even when reporting full cooperation, were "reticent to call it deliberate." This converges with the many subject reports from the present study where unsolicited elaborations of their answers were written alongside the MC items. Often these explanations told of the subject's cooperation, but disallowed that they were purposely trying to "fudge the data" or respond in that manner only because the experimenter expected them to. In other words, it seems likely that subjects can and do give valid reports of their various aspects of awareness (contingency, hypothesis, evaluation) but that a description of their motivation is going beyond their introspective capacities. This seems very reasonable from the perspective of the subject, in that successful problem solving in the conditioning paradigm would seem to depend upon his very conscious awareness of the correct contingency and hypothesis. However, while participating in the experiment, he is not directly concerned with the question of his cooperation or intention, and it is not until the experimenter asks him later that these attributions have any reality at all. At this point they are likely to be greatly influenced by the precise wording of the question, and have an ambiguous status which would not be systematically related to other components of awareness or to situational manipulations.
In this regard, the present findings are congruent with speculations by both Orne and Rosenberg about the degree to which subject motivation is verbalizable to the subject. Orne has recently stated that

It was never my impression, except in rare cases, that the mechanisms by which demand characteristics affect subject's behavior were those of willful or conscious compliance--the subject is not being compliant in any useful sense of that word. Rather he is behaving in ways that, unthinkingly, he perceives as correct or appropriate [1970, p. 225].

And, regarding the relationship between evaluation awareness and motivation to comply, Rosenberg has taken a similar position regarding the ability of subjects to verbalize their cooperation, saying that

From interviewing conducted after data collection in this study and in others I have formed the impressions that subjects will usually obscure from themselves the extent to which they regulate their responding so as to win favorable judgements from the experimenter [1969, p. 296].

While these two statements are based on different reasoning, Orne's suggesting a simple non-verbal process, and Rosenberg, I think, suggesting a more dynamically based process, both are consistent with the present research and with the studies of Bottom and Uleman, regarding the lack of subject reports of MC to be related to other components of awareness in a systematic manner.

The Controversy Regarding Awareness

In the present study, verbal reports of awareness accounted for almost twice as much variance (.401) in the dependent variable of conditioning than did all the manipula-
tions, including their interactions (.238). Certainly within this paradigm, at least, subjects are able to give more insight into their behavior than the experimental psychologist who might depend solely upon the effect of the independent variable manipulations on overt conditioning. This finding supports the phenomenological viewpoint which suggests that we might learn more about behavior simply by asking our subjects about relevant features of their experience.

However, the question of whether such awareness plays a mediational role in the production of conditioning is a more complicated issue, which no experiment, including the present, is able to answer conclusively, although a number of issues emerge from the data which have interesting implications.

A brief review of the controversy will help clarify these. Recall that the extreme cognitive position suggests that subjects approach the conditioning situation with a problem solving set, entertain and discard various hypotheses regarding the contingencies of reinforcement, and decide whether or not to cooperate with the demand characteristics of the experiment. This viewpoint places awareness of the features of the experiment, such as the contingencies and the experimenter's expectations, in a central role of mediating the effects of stimuli on the resulting behavior.

The descriptive behavioristic position, on the other hand, maintains that awareness, even if it is demonstrated,
plays no causal role in the conditioning effect. There are several variations on this position as outlined by Spielberger and DeNike (1966). The first is that awareness is suggested by the interviewing procedures, the second is that it is an artifact of labeling the conditioning effect which the subject notices, and a third suggests that both awareness and conditioning are simultaneous effects of the contingencies, the awareness merely being a verbal variation of the response.

Past research has been designed to differentiate between these explanations, particularly within the verbal-conditioning paradigm where the issue first came into focus when Greenspoon and Taffel concluded that conditioning without awareness took place in their original studies. The verbal-conditioning paradigm is particularly well suited to speak to the third explanation of joint conditioning because the behavioral responses involve a temporal dimension in which the occurrence of awareness can be compared to the occurrence of the conditioned response. When such trend analysis has been performed, what has typically been found is sharp increments in performance following immediately after the inception of awareness, rather than gradual increments in both, which the joint-conditioning model would predict (Spielberger, 1962).

The classical-conditioning paradigm is well suited for speaking to the labeling explanation, since unlike the
verbal-conditioning paradigm, the classical-conditioning paradigm uses a single response on the part of the subject which can be collected before or after the awareness measures. When this has been done, the pre-post factor has repeatedly shown no effect (Bottom, 1972; Weber & Riddell, 1972). This suggests that it is unlikely that the awareness reports are merely labels which subjects give to describe the change in behavior which they notice from the dependent variable.

The notion that awareness is suggested by the interviewing techniques designed to assess it has also received some empirical attention within both paradigms. Within the classical-conditioning one, a comparison of questionnaire formats by Weber and Riddell (1973) elucidated the incidences of awareness assessed by direct questioning versus funnel technique, that is, the employment of general questions followed by specific items to assess awareness. Their results did not support Staats' claim (1969) that the funnel technique produces awareness, in that the presence of the general questions did not enhance the incidence of awareness. These global items, which had been previously used by themselves to define awareness (Staats & Staats, 1957) were shown to be deficient in that more awareness was detected with the specific items, and specifically measured awareness closely predicted actual conditioning. In fact, no evidence for conditioning was found when aware subjects from the specific items were removed from the analysis.
Evidence concerning the verbal-conditioning paradigm also supports the view that awareness is not necessarily an artifact of interviewing technique. Specifically, Klein and Weiner (1966) assessed awareness after a Taffel task with a "Visual Recognition Threshold Test" in which subjects were to read successfully clearer carbon copies of the critical contingencies, and awareness was measured through the number of copies necessary to read the contingency sentence correctly. This method of assessment prevented suggestion, and still, subjects with higher awareness levels conditioned, while low aware subjects did not.

These studies, while not proving the mediational role of awareness, do converge to suggest that the various behavioristic explanations are not supported by empirical evidence from both paradigms.

It is not the author's intention to suggest that awareness is always essential for performance and learning to take place, for common sense generates numerous examples in which habits are maintained without the presence of awareness. Further, with regard to the issue of learning without awareness, Kimble (1962) cites some extreme examples, such as the classical conditioning of intestinal cells to the administration of salt solutions, as well as the conditioning of the eyeblink response, in which learning is not dependent upon the correct verbalization of the CS-UCS relationships. However, it does seem plausible that in the classical conditioning of attitudes paradigm used in this study, awareness plays a central role in
the production of the effect. This evidence comes from the repeated demonstration of no learning without awareness (Page, 1969; Weber & Riddell, 1972, as well as the present data). The intuitive appeal of this rests on the nature of the stimulus situation and conditioned response: a highly unique set of contingencies shaping a written response which certainly involves the use of striated muscles.

With regard to the causal role of awareness, however, the issue is still unresolved. Gross analysis of selected groups of hypothesis aware and unaware subjects, suggests that a knowledge of the experimenter's expectations is necessary for the conditioning effect, as well as for the effect of evaluation cues on resulting behavior. However, the series of partial correlation predictions derived from the Dulaney model did not enable conclusions with regard to the causal effects of these components. This was due largely to the severe problems in measuring motivation to comply. However, predictions from the Dulaney model which did not involve the MC term also failed to be supported. That is, conditioning should have been more highly related to hypothesis awareness than to contingency awareness, which it was not.

One problem with research designed to test this model involves the fact that the components are difficult to measure, and so non-supportive findings are often dismissed as measurement problems, and the legitimacy of the model is retained. However, the present research has shown non-support, even though some of the predictions involve terms which have
been subjected to substantial validation testing. In this context, the present findings suggest that the model may need revision. It is the author's opinion that behavior in the present paradigm is an example of operant, rather than classical conditioning, and that awareness of the contingencies is important for the production of the critical response. However, the Dulaney model has not functioned to support this perspective in the present case, and it may be that the precise role of awareness is much more difficult than the model allows.

Indeed, Page (1973) and others (Doctor, 1971; Greenspoon & Brownstein, 1967; Goldstein et al., 1972) have pointed out it is difficult, if not impossible, to design a study which critically tests the question of whether awareness is necessary for conditioning. What is appropriate, instead, is a series of observations which make either explanation less and less plausible. The present study serves this purpose. Even though the Dulaney model was not supported, subjects who showed no hypothesis awareness did not condition, in line with a series of studies which have reported this finding (Page, 1973; Weber & Riddell, 1973; Hare, 1964). Further, hypothesis awareness served as a critical mediator for the effects of evaluation cues on conditioning. Past studies have shown that manipulations which detract from hypothesis awareness reduce conditioning (Page, 1969) with the use of filler items to make hypothesis awareness more difficult; the use of naive introductory students who are less
familiar with the concept of conditioning). The present study speaks to the opposite effect, that is, enhancing hypothesis awareness through the explicit delineation of contingency information increased conditioning. In addition, Page (1973) has made a similar observation by enhancing hypothesis awareness through instructing the subjects to look for consistent patterns between the nonsense syllables and meaningful words. Thus, the present study fits within a larger body of research which suggests that hypothesis awareness is essential for the production of the conditioning effect in this particular paradigm. While it is not possible to unequivocally rule out the behavioristic explanation, it becomes more and more difficult for the role of awareness to be discounted as such research accumulates.

**Current Status of Subject Roles and Their Implications**

In what manner has the current research contributed to our understanding of subject motivation in the laboratory?

Up to now, both experimental and observational evidence seemed to favor the apprehensive role as the predominant subject motive, and the good subject as merely a facet of apprehension. This resulted from an extensive review of subject-role literature (Weber & Cook, 1972) as well as laboratory studies designed to pit the roles against each other (Sigall, *et al.*, Geller & Endler, 1973). However, lack of adequate manipulation checks and other design problems
showed that these conclusions were unwarranted.

The current study employed extensive manipulation checks and found cooperation and apprehension both operative, and that it is not viable to reduce one role to the operation of the other, as Weber and Cook tentatively suggested. Contingency information clearly enhanced conditioning, although it did not seem to affect evaluation awareness, and over all subjects, contingency and hypothesis awareness showed no relationship to evaluation awareness. The fact that cooperation varied when apprehension was constant makes it inappropriate to explain cooperation solely with recourse to the concept of evaluation apprehension. Even more important, given the complex ways in which situational and person variables interacted in the present study, the notion of "role" is misleading in the first place. Cooperation and resistance are neither characteristics of subjects or of situational demands, but instead a function of the interplay between both.

The present study also helps to resolve a previously posed question regarding the specific way in which evaluation apprehension mediates compliance with demand characteristics. Previous research on the effects of apprehension with regard to the experimenter expectancy effect (EEE) (Duncan, Starkey, Rosenberg, & Finkelstein, 1968) has shown that voice quality cues were effective in producing the effect, except when evaluation awareness was explicitly decreased. When apprehension was enhanced through cues, or left at its normal
level, the EEE was demonstrated. However, as the authors point out, the data did not permit them to distinguish whether the apprehension operated to help the subjects discern the demands, or whether it operated after that to induce motivation to comply with them. While there are striking differences between their Picture Rating paradigm and the classical conditioning one used in the present study, the fact that in the present case evaluation cues did not increase either contingency or hypothesis awareness suggests that the second interpretation may be the most appropriate.

While the effects of the evaluation and hypothesis cues were orthogonal with respect to awareness, the eventual way in which they influence cooperation appeared to be interactive. That is, which "role," cooperation or apprehension, was elicited, depended upon the specific combination of cues in the experimental situation, as well as upon the subjects' need for approval. It was not possible to specify the effects of one cue without determining the levels of the other factors. Hypothesis cues were centrally important, in that they were responsible for the conditioning, and also necessary for the effectiveness of the evaluation cues.

This series of findings suggests that both Orne and Rosenberg are speaking of valid subject concerns, and that the extent to which either is elicited depends upon the particular experimental situation as well as the particular subjects. In this regard, it is less meaningful to ask which "role" is more important, than to specify the specific experi-
mental procedures which would give rise to either role and the specific people in our studies. To the extent that the classical conditioning of attitudes paradigm employs procedures which are not generalizable to other paradigms, conclusions from the present study are limited, in that a key implication from this research is that roles are situationally and person specific. In what way, then, has the current research contributed to the question of how to conduct methodologically sound research in other paradigms which take into account the importance of these very real subject motives?

First, with regard to specific cues used in the present research, the effects of the personality test and prior information communicated by the accomplice suggests that experimenters should take great care to reduce evaluation cues of these sorts. If tests are a feature of the research, they should be administered at a different time and setting than the collection of the experimental data. Experimenters should take care to reduce the possibility of subjects communicating the purpose of the research, even though this is at times difficult. One way to do this would be to debrief subjects all at once, in the manner of the present study. All cues which suggest the hypothesis of the experiment, such as the title, introductory information, and experimental procedures, should be examined for possible sources of hypothesis and evaluation cues. Research in which the dependent variable is collected away from the experimental setting, such as non-reactive measures in the field, or a
change in experiments, would seem to be particularly useful.

At a minimum, experimenters should assess the impact of any demand characteristics by assessing in detail post-experimental awareness, since such reports repeatedly show utility in the exploration of subject cooperation. The present study indicates that it is possible to measure these components of awareness with varying degrees of success. Null findings with the motivation to comply measures from this research, as well as from other studies, indicate that it may be difficult if not impossible to gain much from subjects' reports of compliance. A recent paper by Insko and Cialdini (1971) suggests this may be because compliance in social influence situations is a much more common habit than is the recognition of contingencies and procedures in the experimental laboratory. For this reason, it seems likely that hypothesis and evaluation awareness would be more firmly anchored to verbal behavior than would measures of motivation to comply.

However, the present study has gone further in developing and validating a measure of evaluation awareness than previous research. In this respect it supports Alexander and Knight's (1971) demonstration that subjects act in line with "situated identities," that is, favorable character inferences, but it suggests that a more meaningful way to measure such awareness may be through the use of open-ended essay questions rather than through the use of semantic differential type attributions.
In any case, since reports of awareness showed twice as much relationship to the dependent variable than did all the manipulations and their interactions, it is evident that subject awareness data are useful to the better understanding of behavior in this paradigm. It seems reasonable that this would be the case across a wide number of other paradigms, even though the specific relationships between situational cues and subject cooperation would be expected to change with a change in procedural cues. This research has underscored the view of the human subject as an active problem solver whose behavior is monitored on the basis of various types of demand characteristics. Certainly, as Orne (1969) suggests with his notion of quasi-controls, behavior which is uncorrelated with reports of key hypotheses would be much less likely to be a product of subject cooperation.

Reports of awareness also serve the crucial need for manipulation checks of independent variables. While such checks have been recommended in the past, particularly for investigations of more molar variables such as those within social psychology, the dearth of such checks within the context of subject-role investigations is truly astonishing.

Thus another series of recommendations derived from this research related not so much to the content of the findings, but to the approach taken to the questions. That is, the triangulation of observations on a phenomenon is the essential logic and strategy of construct validity (Cronbach & Meehl, 1955) and upon which all psychological laws will
eventually rest. The investigation of various relationships between data patterns within a single study such as the present would seem to be particularly useful, since differences in laboratories and paradigms can be ruled out as explanations for any incongruencies in findings, providing a more complete test of the processes under investigation. The findings concerning the individual difference measure of Social Desirability is a case in point.

This measure was chosen because it was conceptualized as an indication of a chronic variable, analogous to the acute manipulation of evaluation apprehension by the evaluation cues. Chronic variables indicate natural levels of a construct which exist without manipulation, whereas acute levels are induced through experimental manipulations (see McGuire, 1968). The same disadvantages of each type of variable exist as with the experimental and correlational approaches, that is, the use of chronic variables risks confounding through subject self-selection, and the use of acute variables risks confounding through the artificiality of manipulation. When predictions fail to be supported it could be because of either of these inherent problems.

However, by combining both chronic and acute approaches within the same study, it is possible to maximize information output by converging observations. This was the intent behind the use of the SD measure in concert with the evaluation cue manipulations.
The fact that the SD measure functioned in the hypothesized manner with evaluation cues in the present study is inspiring, although the lack of its power to produce significant simple effects calls for further research. Two suggestions have already been made—trichotimization and running subjects individually.

The SD measure was chosen because it was the closest measure of the evaluation apprehension construct, and it had the most construct validity evidence behind it. However, it would seem to be measuring an extremely general trait, i.e., the propensity to be socially influenced, while a more specific trait might be far more predictive of cooperation in an experiment. In this regard, a measure of subject attitudes about the specific situation, i.e., the psychological experiment, might be far more predictive. Past research with volunteers versus non-volunteers (Rosenthal & Rosnow, 1969) indicates that the degree to which subjects willingly engage in experiments affects their behavior. Other research (Nottingham, 1972; Gustav, 1962; Straits, Wiebeen and Tehophile, 1972) show that prior attitudes of subjects about experiments affect their behavior. Perhaps a measure of pre-experimental attitudes would have functioned more effectively in place of the SD scale.

It is interesting to note that recent attempts by Adair have been directed toward this end. His "Psychological Research Survey" which consists of 52 statements of various attitudes toward experiments, psychologists and science,
was found (over a series of studies) to predict conditioning in the verbal operant paradigm, as well as greater attitude change and conformity (Adair, 1970a, b; 1972; Adair & Fenton, 1972). However, at present, the scale is very heterogeneous in content, and it has not been possible, due to lack of adequate experimental assessment, to decipher whether the incoming attitudes affect motivation to comply with demands, or the original perception of them, as Adair has maintained. Given the preliminary positive findings, it would seem that an appropriate next step in the investigation would be to factor analyze the scale, and relate specific factors to the incidences of awareness and cooperation. In this manner, it would be possible to develop individual difference measures which might more powerfully predict differences in perceptual and cooperative processes among different types of subjects, and which could then be used for other research paradigms.

In a more general sense, the key to understanding the role of subject motivation will ultimately rest on the formulation of adequate theory which would take into account the various known sources of contamination and tie them to a few mediating constructs. Researchers could then design studies which control for the occurrences of these constructs. An interesting start in this direction has recently been made by Rosnow and Aiken (1973).

Using a combination of role theory and McGuire's social influence theory (1968; 1969), they postulate that
subjects, like other socialized human beings, are susceptible to demands created by others. Like the hat-check girl who leaves two quarters in her basket knows (Page, 1973), people don't have to be literally told of these demands. The demands are discerned through various situations, as a result of consistency in the social order. One of these situations occurs within the psychological laboratory, where, as in other social influence situations, people are susceptible to the extent that these demands are communicated and they yield to them. Thus, factors which enhance receptivity, such as the EEE, titles of experiments, curious machinery, etc., and those which affect acquiescence, such as volunteer status, evaluation apprehension, and reactance, are effective in producing laboratory artifacts because they affect these mediators. Researchers can design valid research to the extent that they skirt these mediators entirely, as in non-laboratory research, or manipulate the mediators and show that they do not affect the dependent variable (as in triangulation and replication over irrelevant variables; Campbell & Fiske, 1959; Brunswik, 1947).

Thus, the social psychological features of the laboratory are simply specific cases of the determinants of social behavior in general. To the extent that antecedents of such phenomena as conformity, compliance, ingratiation, reactance, and impression management are linked to the phenomena up to now labeled "demand characteristics" and
"evaluation apprehension," artifacts will be easier to specify and research findings will become generally more interpretable.
BIBLIOGRAPHY


Adair, J. G. Pre-experiment attitudes toward psychology as a determinant of subject behavior. In Methodological problems in research with human subjects. Symposium presented at the meeting of the Canadian Psychological Association, Winnipeg, May, 1970. (b)


Adair, J. G., & Schachter, B. S. To cooperate or to look good?: The subjects' and experimenters' perceptions of each others' intentions. Journal of Experimental Social Psychology, 1972, 8, 74-85.


Hare, R. D. Cognitive factors in transfer of meaning. Psychological Reports, 1964, 15, 199-206.


Milgram, S. Some conditions of obedience and disobedience to authority. *Human Relations*, 1965, 18, 57-76.


Orne, M., & Schiebe, K. The contribution of nondeprivation factors in the production of sensory deprivation effects: The psychology of the "panic button." *Journal of Abnormal and Social Psychology, 1964, 68,* 3-12.


APPENDIX A

INSTRUCTIONS TO SUBJECTS

Personality Test Instructions

Before you is a personality test. Personality tests of this sort can't measure all of a person's personality, of course, but they do tap a specific aspect of psychological functioning in a systematic way. Please be careful and honest when filling this out. Beside each item, place a T if you basically agree with its content, or an F if you basically disagree. For items which are difficult to answer, put the answer which first occurs to you. If you have any questions, please ask the experimenter now.

General Instructions

This is a study of intermodality learning. It involves the association of visual and auditory information, that is, the association of information you receive through your eyes and the information you receive through your ears. The visual information will consist of nonsense syllables which I will flash up on the screen in front of you. The auditory information will be words which I will read as each syllable is flashed.

[Contingency Information spliced in here]:

It will become obvious after only a few of these pairings that two of the syllables flashed on the screen will be special. They will be special because
one of them, YOF [WUH every other condition] is always going to be paired with words of pleasant or positive meaning. Likewise, WUH [YOF every other condition] is always going to be paired with words of unpleasant or negative meaning. The other three syllables are always going to be paired with words of neutral meaning.

There are two tasks which I would like you to carry out. The first is to pay attention to the nonsense syllables which are flashed. The second is to repeat the word which I say after I say it in a loud clear voice. For example, the first pair will consist of the nonsense syllable LAJ which will be flashed on the screen while I say the word "with." Your job is to pay attention to the syllable on the screen, and at the same time to clearly repeat the word "with" together as a group, after I say it. There are eighty pairings; this will take about five minutes. If you have any questions, please ask the experimenter now.

**Dependent Variable Instructions**

Please turn over the syllable booklet in front of you and clearly print your full name in the top right corner. In this booklet there are a number of nonsense syllables. Your task consists of two parts: first, rate each syllable according to how pleasant or unpleasant you feel it is. A mark in the middle of the scale indicates no feeling or a completely neutral feeling. Please put your marks between the colons, rather than on the colons.
Second, check each syllable as to whether you think it was presented in the part of the experiment just completed. Even if you think a syllable wasn't presented, you should rate how pleasant it is anyway, so that when done, you will have checks and pleasantness ratings for each syllable. If you have any questions, please ask the experimenter now.

[Accomplice asked the question in the Accomplice conditions here]

**Post-Experimental Questionnaire Instructions**

Even though the experimental aspects of this study are now over, a very important kind of information still remains to be gathered, and that is your experience of the experiment itself. This is, in fact, one of the most important aspects of this research. Actually, this research is probably different from others that you might participate in, in that here, it is extremely important that we know how you, as a subject, interpreted the situation. Most experimenters merely gather the measures, and send you on your way, never giving you a chance to tell what you think was going on. So would you please be as conscientious and as truthful as possible when recalling your experience in the experiment itself.

There are some instructions which I would like to go over with you on this questionnaire. Please follow along with me. Number one: please answer the items in the order in which they appear. Do not go back and change answers. If a new idea occurs to you which would lead you
to want to change one of your answers, jot the new thought down as soon as you think of it. Number two: Make sure you understand each question before you answer it. If something is not clear, be sure to ask the experimenter. Number three: please be as honest and as conscientious as you can. There is just one more thing that I would like to add. Most of the time, subjects come into an experiment with a role they play. They are supposed to be passive and naive. However, I want you to respond to this questionnaire as if you are the expert which you really are. For when it comes to your experience, I am naive, and you know everything. From this standpoint, I would like you to be active, rather than passive. If an idea occurs to you which is only indirectly related to the item you are answering, include it. Remember, any and all information which you can tell me about how you went through this experiment will really help me in this research. Thank you.
APPENDIX B

MARLOW-CROWNE SOCIAL DESIRABILITY SCALE

1. Before voting I thoroughly investigate the qualifications of all the candidates.
2. I never hesitate to go out of my way to help someone in trouble.
3. It is sometimes hard for me to go on with my work if I am not encouraged.
4. I have never intensely disliked anyone.
5. On occasion I have had doubts about my ability to succeed in life.
6. I sometimes feel resentful when I don't get my way.
7. I am always careful about my manner of dress.
8. My table manners at home are as good as when I eat out in a restaurant.
9. If I could get into a movie without paying and be sure I was not seen I would probably do it.
10. On a few occasions, I have given up doing something because I thought too little of my ability.
11. I like to gossip at times.
12. There have been times when I felt like rebelling against people in authority even though I knew they were right.
13. No matter who I'm talking to, I am always a good listener.
14. I can remember "playing sick" to get out of something.
15. There have been occasions when I took advantage of someone.

16. I'm always willing to admit it when I make a mistake.

17. I always try to practice what I preach.

18. I don't find it particularly difficult to get along with loud mouthed, obnoxious people.

19. I sometimes try to get even rather than forgive and forget.

20. When I don't know something I don't at all mind admitting it.

21. I am always courteous, even to people who are disagreeable.

22. At times I have really insisted on having things my own way.

23. There have been occasions when I felt like smashing things.

24. I would never think of letting someone else be punished for my wrong doings.

25. I never resent being asked to return a favor.

26. I never have been irked when people expressed ideas very different from my own.

27. I never make a long trip without checking the safety of my car.

28. There have been times when I was quite jealous of the good fortune of others.

29. I have almost never felt the urge to tell
someone off.

_____30. I am sometimes irritated by people who ask favors of me.

_____31. I have never felt that I was punished without cause.

_____32. I sometimes think when people have a misfortune they only got what they deserved.

_____33. I have never deliberately said something to hurt someone's feelings.

_____34. Sometimes I take a disappointment so keenly that I can't put it out of my head.
APPENDIX C

DEPENDENT VARIABLE SYLLABLE BOOKLETS*

G I N

pleasant : unpleasant

This syllable was ______ presented before

was not ______

---

V E C

pleasant : unpleasant

This syllable was ______ presented before

was not ______

* Each syllable appeared on a separate page, as indicated by the double lines.
F I W

This syllable was _____ presented before
was not _____

H U E

This syllable was _____ presented before
was not _____

L A J

This syllable was _____ presented before
was not _____
YOG

pleasant : ______:________:________:________:________:__ unpleasant

This syllable was _____ presented before
was not ______

---------------

GAH

pleasant : ______:________:________:________:________:__ unpleasant

This syllable was _____ presented before
was not ______

---------------

XEH

pleasant : ______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:______:___
This syllable was ______ presented before
was not _______.

This syllable was ______ presented before
was not _______.

This syllable was ______ presented before
was not _______.
QUG

pleasant :____:____:____:____:____:____:____:____:____: unpleasant

This syllable was _____ presented before
was not _____


WUH

pleasant :____:____:____:____:____:____:____:____:____: unpleasant

This syllable was _____ presented before
was not _____


XAD

pleasant :____:____:____:____:____:____:____:____:____: unpleasant

This syllable was _____ presented before
was not _____
HOF

pleasant :____:____:____:____:____:____:____:____: unpleasant

This syllable was _____ presented before

was not _____


VAF

pleasant :____:____:____:____:____:____:____:____: unpleasant

This syllable was _____ presented before

was not _____
APPENDIX D
CONFORMITY TEST

1. I often find myself going along with a group even though I privately do not agree with them.
2. Before I vote on an issue, I like to consult my friends and people I respect to make sure my opinions are not too contrary to theirs.
3. I dread door-to-door salesmen because I'm gullible and can't say "no".
4. I am more cooperative and persuadable than most people I know.
5. I find it very uncomfortable to be in disagreement with anyone.
6. I am easily influenced.
7. I would rather change my mind on an issue than be the only person in the group to take a different stand.
8. I often worry about strangers looking at me and laughing at something I am wearing.
9. I try to keep up with the latest fashions.
10. It's important to be in the "in crowd".
11. I worry about having offended people accidentally.
12. I often find that I am easily persuaded by an argument which I later find is not well founded.
13. I have had some people tell me that I am too gullible.
14. I always try to cooperate with other people's requests.

15. I think it's important not to "rock the boat."

16. The world would be a lot better off without the irrational emphasis on individualism.

17. I like to follow instructions and to do what is expected of me.

18. I think cooperation is more important than individualism.

19. I suppose you could say that I am a conformist.

20. I like to tell other people the things they like to hear, even if I privately do not agree with what I am saying.
Please follow these instructions very carefully when filling this out--remember, this is a most essential part of this research.

1. Answer the items in the order in which they appear. Do NOT go back and change answers. If a new idea occurs to you which would lead you to want to change one of your answers, jot the new thought down as soon as you think of it.

2. Make sure you understand each question before you answer it. If something is not clear, be sure to ask the experimenter.

3. Please be as honest and as conscientious as you can.

1. How valuable do you think this experiment is scientifically?

Circle one: 1. Very valuable
2. Valuable
3. Somewhat valuable
4. Minimally valuable
5. Completely valueless
6. Don't have any idea

Why?

2. How interesting did you find this experiment?

Circle one: 1. Very interesting
2. Interesting
3. Somewhat interesting
4. Minimally interesting
5. Completely boring

3. Would you say that how a person rated the syllables is more a function of the situation or of his particular personality?

_____Situation
_____Personality
_____Both equally important
4. What was the purpose of the personality test given at the beginning of the experiment?

5. What specific aspect of personality did the test measure?
6. What nonsense syllable was always or usually paired with spoken words of pleasant or positive meaning?

7. What nonsense syllable was always or usually paired with spoken words of unpleasant or negative meaning?
8. How do you think the experimenter expected people to rate YOF?

Pleasant:____:____:____:____:____:____:____: Unpleasant

9. How do you think the experimenter expected people to rate WUH?

Pleasant:____:____:____:____:____:____:____: Unpleasant
10. If a person rated the nonsense syllables WUH and YOF according to the pleasant or unpleasant words they were paired with, would this indicate anything about his personality or intelligence? (Check one)
   - A. No, I do not think that how a person rated WUH and YOF and WUH in this experiment would have anything at all to do with either his personality or intelligence.
   - B. There may or may not be some relationship between the way a person rated WUH and YOF in this experiment and his personality or intelligence, but if there is, I do not know what it would be.
   - C. Yes, I think there is some relationship between a person's rating of YOF and WUH in this experiment and his personality or intelligence, but I don't know what it is.
   - D. Yes, I think there is some relationship between a person's ratings of YOF and WUH in this experiment and his personality or intelligence, and I might know what it is.
   - E. Yes, I think there is some relationship between a person's ratings of YOF and WUH in this experiment and his personality or intelligence, and I think I know what it is.
   - F. Yes, I think there is some relationship between a person's ratings of YOF and WUH in this experiment and his personality or intelligence, and I definitely know what it is.

11. If a person rated the nonsense syllables WUH and YOF according to the pleasant or unpleasant words they were paired with, what would this indicate about his personality or intelligence?

12. If a person did not rate the nonsense syllables WUH and YOF according to the pleasant or unpleasant words they were paired with, what would this indicate about his personality or intelligence?
13. On the following dimensions, give your impressions of the typical person who would rate the syllables YOF and WUH according to the positive or negative words they were paired with in this experiment:

Normal:________:________:________:________:________:________:________:Abnormal
Intelligent:________:________:________:________:________:________:________:Stupid
Individualistic:________:________:________:________:________:________:________:Conforming
Cooperative:________:________:________:________:________:________:________:Uncooperative
Not Gullible:________:________:________:________:________:________:________:Gullible
Pleasant:________:________:________:________:________:________:________:Unpleasant
Likeable:________:________:________:________:________:________:________:Disagreeable
Sincere:________:________:________:________:________:________:________:Insincere
Flexible:________:________:________:________:________:________:________:Rigid
Trustworthy:________:________:________:________:________:________:________:Untrustworthy
Well Adjusted:________:________:________:________:________:________:________:Maladjusted
Open Minded:________:________:________:________:________:________:________:Closed Minded

14. During the experiment, did you want to rate the syllables the way you thought the experimenter expected you to?

Wanted to very much:________:________:________:________:________:________:________:did not
Somewhat did Did not care Somewhat did not
15. In this type of study, some people are influenced by knowing the experimenter's expectations to report perhaps a little more feeling about the syllables than they actually experienced in order to "help" the hypothesis along. Others react by trying to act naturally and not "help" the hypothesis along. Still others react in a negativistic way and actually try to resist rating the syllables according to the hypothesis, sometimes even doing the opposite of what they think the experimenter is looking for. With this in mind, and its important to be honest at this point, how much did you try to help the hypothesis along by rating the syllables according to the way you thought you were supposed to?

<table>
<thead>
<tr>
<th>Tried to</th>
<th>Tried to do</th>
</tr>
</thead>
<tbody>
<tr>
<td>very much: _____</td>
<td>: _____: _____: _____: _____: the opposite: _____: _____: _____: _____: the opposite</td>
</tr>
<tr>
<td>Acted naturally</td>
<td></td>
</tr>
</tbody>
</table>

16. To what extent did you resist rating the syllables according to the experimenter's expectations in order to avoid looking gullible or conforming?

<table>
<thead>
<tr>
<th>Resisted very much: _____: _____: _____: _____: _____: _____: _____: _____: _____: _____: _____: at all</th>
</tr>
</thead>
</table>

17. Please write down any other comments which you feel might help us to understand your reaction to this experiment.