

This dissertation has been
microfilmed exactly as received

66-10,189

PAGE, Monte Merle, 1935-

THE EFFECT OF REWARD AND PUNISHMENT ON
FIGURE-GROUND PERCEPTION AS A FUNCTION
OF DEMAND CHARACTERISTICS AND PSYCHO-
LOGICAL SOPHISTICATION OF SUBJECTS.

The University of Oklahoma, Ph.D., 1966
Psychology, experimental

University Microfilms, Inc., Ann Arbor, Michigan

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE EFFECT OF REWARD AND PUNISHMENT ON FIGURE-GROUND PERCEPTION
AS A FUNCTION OF DEMAND CHARACTERISTICS AND
PSYCHOLOGICAL SOPHISTICATION OF SUBJECTS

A DISSERTATION
SUBMITTED TO THE GRADUATE FACULTY
in partial fulfillment of the requirements for the
degree of
DOCTOR OF PHILOSOPHY

BY

MONTE MERLE PAGE

Norman, Oklahoma

1966

THE EFFECT OF REWARD AND PUNISHMENT ON FIGURE-GROUND PERCEPTION
AS A FUNCTION OF DEMAND CHARACTERISTICS AND
PSYCHOLOGICAL SOPHISTICATION OF SUBJECTS

APPROVED BY

AE Duhell

William H. Tinsdale

Adams

William Tins

C. F. Shuman

DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

The author wishes to thank the members of the dissertation committee, Dr. Alfred Glixman, Dr. Richard Terry, Dr. William Trousdale, and Dr. William Terris for their time and interest. To Dr. Arnold E. Dahlke, my committee chairman, I wish to express special appreciation. Because of his able direction and keen interest, the progress of this research was relatively smooth.

Special mention also goes to Dr. J. D. Palmer and Dr. Richard Terry of the Systems Research Center for providing a room in which to collect data.

Thanks also goes to Clinton Goff, George Letchworth and David Paskewitz, graduate students at the University of Oklahoma, for their assistance.

To my wife, Joyce, goes my deepest appreciation for the typing of this dissertation and for the many other things she has done to help make my years in graduate school possible.

TABLE OF CONTENTS

	Page
LIST OF TABLES	v
LIST OF FIGURES.....	vii
Chapter	
① I. INTRODUCTION.....	1
II. PROBLEM.....	21
III. METHOD.....	27
IV. RESULTS.....	42
V. DISCUSSION.....	71
VI. SUMMARY.....	83
REFERENCES.....	85
APPENDIX A	88

LIST OF TABLES

Table	Page
1. Summary of Characteristics of the Subject Population....	30
2. Mean Percentage of Rewarded Responses for Five Groups in the Experiment.....	46
3. Analysis of Variance for the Data in Table 2.....	46
4. Mean Percentage of Rewarded Responses for Five Groups in the Experiment, Breaking by Reward Order.....	48
5. Frequencies of Rewarded Response Scores Above and Below 50 percent for Aware and Unaware Subjects.....	50
6. Frequencies of Aware Cooperators and Noncooperators Who Score Above and Below 50 Percent Rewarded Responses.....	53
7. Mean Percentage of Rewarded and Punished Test Responses Scored Separately for Each of the Four Figures.....	55
8. Summary of the Analysis of Variance for Awareness, Subject Population, Reward Order, and Figures Within Order for the 77 Subjects Who Could be Unambiguously Classified on Awareness.....	56
9. Frequencies of Subjects Scoring Above and Below 50 Percent Rewarded Responses in the Two Reward Orders.....	59
10. Frequencies of Aware and Unaware Subjects in the Two Reward Orders.....	59
11. Frequencies of Cooperating and Noncooperating Aware Subjects in the Two Reward Orders.....	60
12. Frequency Distributions of Subjects Ratings of Certainty That Two Profiles Fit on the Test and Which Two Profiles Won Money.....	63
13. Frequency of 39 Aware and 38 Unaware Subjects Classified According to Relative Certainty or Uncertainty That Figures Fit.....	65

Table	Page
14. Frequency of 39 Aware and 38 Unaware Subjects Classified According to Relative Certainty or Uncertainty of the Two Profiles Which Always Won Money.....	66
15. Frequency of 39 Aware and 38 Unaware Subjects Classified According to Whether They Actually Saw Two Profiles Together on the Test a Few or Many Times.....	66
16. Frequency of 96 Subjects Who Scored Above and Below 50 Percent Rewarded Responses Classified According to Whether or not They Expected to Win More Money on the Test.....	68
17. Frequency of 39 Aware and 38 Unaware Subjects Classified According to Expectancy of Winning More Money on the Test.....	68
18. Frequency of 29 Cooperating and 10 Noncooperating Awares Classified According to Expectancy of Winning More Money on the Test.....	69

LIST OF FIGURES

Figure	Page
1. Examples of two of the four profiles presented in the training series by Schafer and Murphy (1943).....	4
2. Example of one of the two reversible figures used by Schafer and Murphy (1943) in the post-training series...	4
3. The four individual profile figures used in the training series.....	33
4. The two composite ambiguous profile figures used in the test series.....	34
5. The two nonreversible set breaking profile figures used in the test series.....	35
6. Frequency distributions for (a) the 48 unaware and non-cooperating subjects and (b) the 29 aware cooperating subjects.....	52

THE EFFECT OF REWARD AND PUNISHMENT ON FIGURE-GROUND PERCEPTION
AS A FUNCTION OF DEMAND CHARACTERISTICS AND
PSYCHOLOGICAL SOPHISTICATION OF SUBJECTS

CHAPTER I

INTRODUCTION

Schafer and Murphy (1943) attempted to experimentally demonstrate that a history of reward and punishment would modify perception, particularly in an ambiguous situation. On several training trials, one aspect of a reversible figure, presented tachistoscopically, was contiguously paired with the winning of a small amount of money (reward) while the other aspect was paired with the loss of money (punishment). In a test situation, when the complete ambiguous or reversible figure was presented, subjects reported seeing the rewarded aspect of the figure more often than the punished aspect. This result was taken as strong evidence for the hypothesis that reward and punishment can determine what is seen as figure and what is seen as ground.

In more recent times, investigators of verbal conditioning have been concerned with the problem of subjects becoming aware of the purpose of the experiment and this in turn eliciting cooperation (Dulany, 1961; Levin, 1961; Spielberger, 1962; Farber, 1963; Page, 1964). Findings of learning and awareness experiments strongly suggest that operant verbal

conditioning such as originally reported by Greenspoon (1955) and the classical conditioning of meaning reported by Staats and Staats (1957) occurs only in subjects who can correctly verbalize the purpose of the experiment. So far, these studies have been limited to verbal conditioning. However, the same problem could exist in other experiments, particularly when deception is involved.

Orne (1962) has approached the same problem of subjects aware nesses in experiments as a part of a more general concern with the social psychology of the psychological experiment. He has suggested that what a subject perceives as the purpose of the experiment be called the demand characteristics of the experiment. According to Orne, such demand characteristics can be very important in mediating the behavior observed in psychological experiments. Demand characteristics artifacts place severe limitations upon the generalizability of data obtained in certain artificial laboratory situations.

In view of the substantial empirical literature which is accumulating with regard to the importance of awareness in verbal conditioning and the contemporary concern with the social psychology of experiments, it seemed appropriate to re-examine the Schafer-Murphy effect. While Schafer and Murphy's experiment was concerned with perceptual learning, their procedures resemble those used by Staats and Staats. If the Schafer-Murphy effect were also found to be an artifact of subjects becoming aware of the purpose of the experiment, this would represent an extension of the problem into an area where investigators have not previously considered the possibility of awareness artifacts.

The present dissertation presents the results of an experimental study of the Schafer-Murphy effect in terms of subjects awarenenses or

the demand characteristics of the experiment. The general hypothesis is that the Schafer-Murphy effect is an artifact of demand characteristics. Specific predictions derived from this hypothesis will be spelled out later. If this hypothesis were confirmed, it would be of particular interest because the Schafer-Murphy experiment is a classic study that appears in many general and social psychology textbooks.

The "Autism" And Figure-Ground Experiment

In the 1940's, Gardner Murphy directed a series of studies designed to experimentally demonstrate his theory of "autism" (Levin, Chein, & Murphy, 1942; Proshansky & Murphy, 1942; Schafer & Murphy, 1943). Perceptual "autism" was defined as the organization of cognitive processes in the direction of need satisfaction. A major hypothesis was that a history of reward and punishment would influence perception, particularly in an ambiguous situation.

Schafer and Murphy (1943) gave a small amount of money (two and four cents) as rewards and took money away as punishment. Four training stimuli were presented 25 times each in a tachistoscope at a 1/3 second exposure speed. They consisted on contour lines drawn on the flat side of a half circle so that they appeared as faces or profiles (see Figure 1). These profiles were actually pairs which shared the same contour line when they were fit together in a complete circle (see Figure 2). For some of the subjects, the right half of one reversible contour was rewarded (won money) and the left half of the other (N-D order), while the other halves were punished (lost money). The other subjects received reward and punishment in the opposite order (B-C order), to control for position and contour effects. After 100 training trials (25 for each face) of consistent

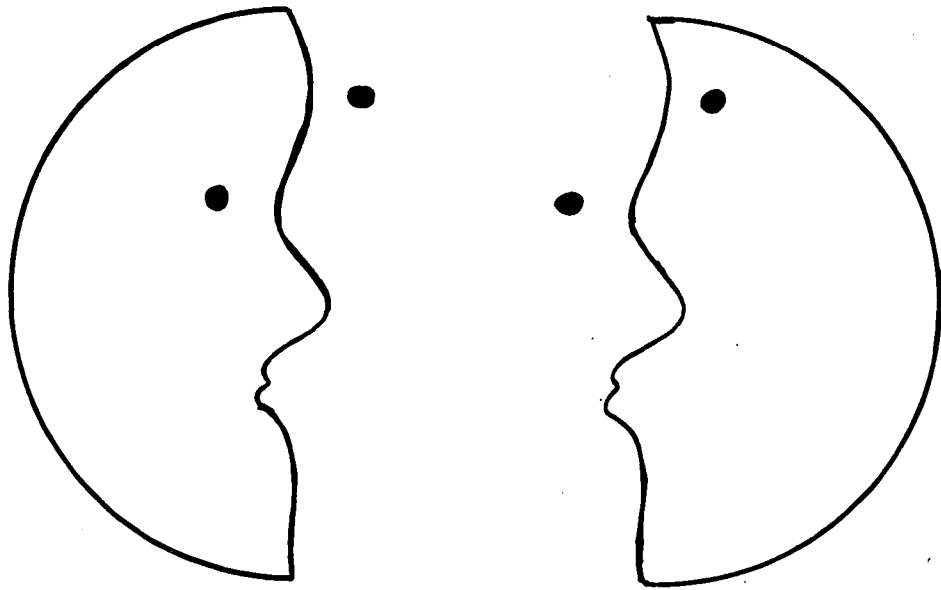


Figure 1. Examples of two of the four profiles presented in the training series by Schafer and Murphy (1943).

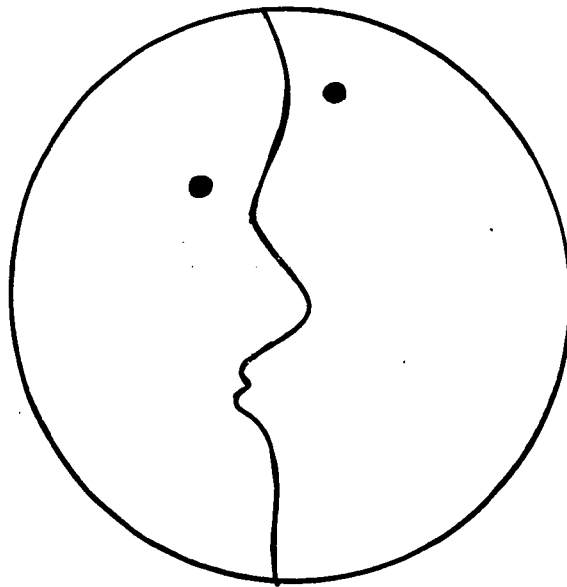


Figure 2. Example of one of the two reversible figures used by Schafer and Murphy (1943) in the post-training series.

reward for two faces and consistent punishment for the other two faces, a test series was initiated. During the test, the pairs of stimuli which shared the same contour line were presented together in a circle. The hypothesis was that subjects would see more rewarded profiles than punished ones in the ambiguous situation. No rewards or punishments were given during testing. The subjects were told they would now see the faces inside a circle, and they were to report what they saw. A total of 32 test trials were given. Between test trials, a nonreversible contour in a circle facing in the opposite direction from the subject's last response was presented. This was to break up the formation of directional sets in perceiving the reversible figures. For the first 16 trials, the hypothesis was strongly supported. About 80 percent of the test responses were to the rewarded figures and 20 percent to the punished figures. After 16 test trials, the response broke down and the authors invoke the concepts of extinction and learning of new directional sets to account for this. Though Schafer and Murphy used only five subjects, there appeared to be striking evidence for an effect of reward and punishment on subsequent figure-ground perception.

Rock and Fleck (1950) attempted to replicate the Schafer-Murphy experiment, using high school students as subjects. They used a projector type tachistoscope instead of the Whipple tachistoscope used in the earlier study. Considerable individual differences were found, but these tended to cancel out to produce an overall insignificant result. They did introduce a post-experimental interview which was very suggestive. Many of their subjects reported seeing both faces in the ambiguous test series, but responded with only one name.

Subjects in the Rock and Fleck study made many more incorrect identifications in the test series than did Schafer and Murphy's subjects. This suggested to Jackson (1954) that there was something different between the stimuli of the two experiments. He noted that Rock and Fleck had used a projector tachistoscope which projected a large image on a screen while Schafer and Murphy had used a tachistoscope which cast a small image and had a constantly lighted field. Consequently he felt that a replication using both kinds of apparatus was appropriate. Using a tachistoscope similar to Schafer and Murphy's, he replicated their results. Using a projector tachistoscope, he obtained results that were statistically reliable, but less pronounced. Like Rock and Fleck, he obtained a high proportion of incorrect identifications. Jackson concluded that Rock and Fleck had failed to replicate because their apparatus made the stimuli too difficult to see and not because the Schafer-Murphy effect could not be replicated.

Jackson introduced a modification of the reward and punishment procedure, but apparently this did not affect the results. In the Schafer-Murphy study, rewards and punishments were ostensibly contingent upon the subjects guessing certain numbers while, actually, two profiles were always rewarded and two were always punished. In the Jackson study, the guessing was left out and the rewards and punishments were given directly. The Jackson procedure has been predominantly used in more recent work on the Schafer-Murphy effect.

Jackson presented individual data for his 12 subjects. It is interesting to note that only four of the subjects had high proportions of rewarded responses to both of the reversible test stimuli; two others had high proportions of rewarded responses for one of the test stimuli, but

about equal proportions of rewarded and punished responses to the other test stimuli. It could be argued that only four of the subjects were motivated by the money and therefore were the only ones whose perception was modified by the training. This argument does not hold, however, for the two who learned the response for only one of the two possible ambiguous stimuli because rewards and punishments were equal for both sets of stimuli. If some cognitive process, such as awareness of the experimenter's hypothesis, were operating in this situation, it is reasonable that subjects aware of one contingency would respond to that contingency. One subject responded maximally in the opposite direction from the hypothesis for one of the pairs. This may have been resistance behavior similar to that found in verbal conditioning. The other five subjects, as a group, balance out at about equal responses for the rewarded and punished figures. Three of these five had a very high proportion of mistaken identifications in the test series (about 50 percent). It is probable that these subjects could not recognize the stimuli during training as well as the others. If so, then it is less likely that they would have become aware.

Jackson's data fits very well with what would be predicted if the same mechanisms were working in his experiment as those to be described later in the section on verbal conditioning and awareness, especially if Jackson's subjects could often see both faces in the test series as suggested by Smith and Hochberg's (1954) second experiment.

Smith and Hochberg (1954) replicated the Schafer-Murphy effect using solid figures rather than outlines and using electric shock as punishment. They found a significant preference for the non-shocked figures in the test series, but considerable individual variation. Their apparatus resembled that used by Rock and Fleck (1950). A second

experiment attempted to relate the phenomenon to subception and perceptual defense. However, it was found that at 1/3 second exposure, both figures could often be identified when the subject was given the set to do so. They concluded that explanations that did not invoke prerecognition processes such as subception were more adequate. It is interesting to note that both figures could often be perceived. Such a finding supports a demand characteristics interpretation because it suggests that subjects may have often seen both figures but chose to report the rewarded figure because of demand characteristics. A demonstration of autism would require subjects to see only one figure.

The Schafer-Murphy design was used as the prototype for an extensive research program on perceptual learning at the Menninger Foundation in the late 1950's. Snyder and Snyder (1956) modified the Schafer-Murphy design in order to study the effects of monetary reward and punishment on the auditory modality. When a rewarded voice and a punished voice were presented at the same time, subjects tended to attend more to the rewarded voice. A brief post-experimental interview was conducted, though what was asked was not reported. The authors indicate that seven of their 41 subjects reported awareness. The importance of this is dismissed, however, because all of the between groups variance is not accounted for by these seven subjects. The present discussion of verbal conditioning will point out that brief interviews typically do not detect all of the aware subjects and this, in turn, could account for Snyder and Snyder's results.

Two studies used the Schafer-Murphy design to influence the perception of Figure-ground through the tactual modality (Sommer & Ayllon, 1956; Ayllon & Sommer, 1956) with significant results. Several other Menninger Foundation studies used children as subjects. Earlier studies

(Solley & Sommer, 1957; Solley & Engel, 1960) found strong autism effects using the Schafer-Murphy profiles and also some other ambiguous figures. A later study (Santos & Garvin, 1962) done by different experimenters, but also from the Menninger Foundation, failed to obtain the effect. These authors are skeptical about the generalizability of the Schafer-Murphy effect and suggest that certain uncontrolled variables, such as the relationship between experimenter and subject may be important. However, their failure to obtain the effect is inconclusive because of several significant departures from previous methodology.

A study by Solley and Santos (1958) used operant verbal reinforcement of a reversible Necker Cube. During 256 trials, a Necker Cube improved to favor a left to right perception was presented 101 times and a cube biased for right to left perception was presented 101 times. One of the improved cubes received 70 percent positive verbal reinforcement ("good" or "fine") while the other received only 30 percent. A balanced or reversible test cube was presented 54 times, interspersed among the training trials. For most subjects, there developed a preference for perceiving the reversible cube in the most often reinforced direction. Some subjects developed no preference at all. The interview for awareness was apparently not stressed as it is dismissed in a sentence. The authors make one statement that is interesting from a demand characteristic viewpoint:

It should be pointed out again that three subjects were not included because they reached a point in conditioning where they reported the reinforced aspect of the cube even when being shown the non-reinforced improved opposite!

These subjects were eliminated from the analysis because their behavior did not make sense from the experimenter's point of view. But, from the point of view of the subjects doing what the demand characteristics of the situation call for, their behavior makes a good deal of sense.

The Solley and Santos (1958) experiment, along with a discussion of it by Santos and Murphy (1960), suggests that the thinking of the Menninger Foundation group was very similar to that found in early verbal conditioning studies. Reinforcement or punishment is the key to modifying behavior even when the subject is not aware of the contingency between reinforcement and his behavior. Sheer contiguity of a pleasant or unpleasant affect (produced by reward and punishment) with a perceptual stimulus, modifies the figure-ground properties of that stimulus when it is later presented in an ambiguous form. This happens regardless of what the reward might mean to the subject (i.e. I get rewarded for guessing the correct number), so long as it is contiguous with the stimulus. The subject is apparently viewed as a passive agent. He may be deceived and given incomplete information without any of it making much difference as long as the reward is presented at the same time as the perceptual stimulus. What the subject is aware of is of little importance, therefore a meager effort, if any, is made to find out about awareness.

In light of the learning and awareness findings to which we now turn, subjects' awareness may be extremely important in the Schafer-Murphy situation. A careful and extensive effort should be made to find out what these awarenesses might have been because it is likely that these awarenesses can account for most, if not all, of the systematic variance in the Schafer-Murphy experiment.

Verbal Conditioning And Awareness

Operant verbal conditioning has been attempted using several different procedures. The most often used procedures are those of Greenspoon (1955), which involves subjects saying words or numbers over

an extended time while being "reinforced" for saying words of a certain class by the experimenter saying "good" or "Um hum," and Taffel (1955), which involves subjects constructing a series of sentences using verbs printed on stimulus cards and being reinforced for beginning their sentences with certain pronouns. Subjects in these situations typically emitted more responses in the reinforced category than did control subjects who received no verbal reinforcement. Subjects who verbalized awareness in a brief post-experimental interview were usually eliminated from the analysis so that the results could be attributed to an S-R bond independent of higher cognitive processes.

Tatz (1960) was concerned that Dollard and Miller (1950) had interpreted Greenspoon's (dissertation in 1950, published 1955) results as supporting Thorndike's hypothesis of the direct action of rewards (i.e. law of effect). He felt that this interpretation depended upon the assumption that "awareness" was absent. Previous investigators had taken into account only subjects who could specifically state the experimenter's system for allotting reinforcement. According to Tatz, the problem was more complex: Earlier investigators had not taken into account "partially correct response systems," which could mediate a level of performance higher than chance even though they did not correspond exactly to the experimenter's system for allotting reinforcements. He obtained awareness protocols post-experimentally. Subjects were then classified into one of four groups, (1) stated the system, (2) stated a successful system, (3) stated a partially successful system, and (4) stated no effective system. When the correlated response hypotheses as well as correct ones were taken into account, no evidence was found for verbal conditioning in truly unaware subjects.

Dulany (1961, 1962) felt that verbal conditioning had attracted attention only because it apparently demonstrated learning without awareness in some sense. He was unsurprised to learn that one can influence the actions of another cooperative human being by informing him that his behavior is judged correct or acceptable, but the claim that this occurred outside of awareness could have alternative explanations. It is possible, he said, that the operant conditioning effect was mediated by some kind of verbal control, i.e. behavioral hypotheses and self instructional sets.

Dulany replicated Greenspoon's experiment, except that he stopped subjects after each 55 responses and asked: "What do you think the experiment is all about?" After the experiment he asked further questions. Not one subject stated that plural nouns were correct (Greenspoon's criterion for eliminating awares). There was evidence that the experimental group said more plural nouns than did the controls. However, of his 43 experimental subjects, 11 said that when the experimenter said, "Um hum" it meant they were to associate in the same category. Most of them also said that when he stopped saying "Um hum" it meant they were to change categories. Subjects (N=32) who did not hold these behavioral hypotheses showed no conditioning effect at all. Dulany then did a word association experiment which showed a strong tendency for subjects instructed to associate in the same category to associate plural words with plural stimuli. Therefore, if subjects acted upon the behavioral hypothesis they reported in his first experiment, they should have increased their output of plural nouns.

Dulany suggests that "a correct behavioral hypothesis" be used as a theoretical term for awareness that relates to response selection

in experiments. Further, he suggests that "correlated hypothesis" be used for "incorrect" awareness that correlates with what the experimenter calls correct. In his experiments he has found no evidence for verbal conditioning in subjects not having either correct or correlated response hypotheses.

Farber (1963) also reports difficulty in finding evidence for verbal conditioning in unaware subjects. In an article entitled, "The Things People Say to Themselves," he reports a study where aware subjects are classified as conforming or nonconforming. Conforming aware showed high conditioning, nonconforming awares showed less conditioning and unawares showed no conditioning. Farber says that one might reasonably argue that subjects who learned, became aware because they learned, but that it would be pushing credulity too far to suppose that subjects statements concerning conformity did not indicate planned action. As an example of what he meant by planned action, he lists several verbal reports given by his subjects. One of them was: "Everytime I said 'you', he said 'good' so naturally I said 'you' most of the time, but now and then I said something else to break the monotony."

Levin (1961) and Spielberger (1962) have emphasized the differences in brief and extended post-experimental interviews in detecting aware subjects. They have found that brief and unstructured interviews frequently do not elicit reports of awareness for subjects while more structured and specific questions do. Their findings strongly support the contention that undetected aware subjects who respond at a high rate, account for all of the systematic variance in operant-verbal conditioning. Also, previous studies did not initiate awareness interviews until after extinction procedures. This could have caused aware subjects to discount

their earlier response hypotheses before they were asked concerning their awareness.

Recently, Page (1964) and Letchworth (1965) have extended the implications of the awareness findings in operant verbal conditioning to include the classical conditioning of meaning experiments introduced by Staats and Staats (1957, 1958). Using the extended post-experimental interview technique, they have found that results leading to the inference of verbal classical conditioning of meaning appear also to be an artifact of subjects becoming aware of the purpose of the experiment.

Staats and Staats (1957) introduced their experiment to the subjects as a study of the simultaneous learning of visually and orally presented words. First, there was a conditioning series where two words, connotatively neutral on Osgood's evaluative dimension (Osgood, Suci, & Tannenbaum, 1957), were each presented visually 18 times. One word was always paired with orally presented words having a heavy loading of pleasant evaluative meaning while the other was always paired with unpleasant words. To make the deception more plausible, four other visual words were also presented randomly 18 times each; but these were paired with spoken words having no consistent connotative meaning. A seven-point semantic differential scale of pleasant-unpleasant was introduced after the conditioning procedure to test for the acquisition of the conditioned evaluative response to the two experimental words. The stated purpose of this was "to see if how we feel about words affects how we learn them" (Staats & Staats, 1957). The deceptions were then carried through by giving a learning test. Finally, subjects were asked to write down what they thought about the experiment, especially its purpose. Subjects who stated the correct relationship between the experimental

words and the loaded words were considered "aware" and were eliminated from the analysis. The remaining data revealed a significant difference between the mean semantic differential ratings of the two previously neutral experimental words. No such change took place for the four control words. This effect was attributed to conditioning of meaning without awareness.

In replications focusing on the assessment of awareness in the above studies (Page, 1964; Letchworth, 1965), behavior similar to that reported in operant verbal conditioning was found. Some subjects apparently saw through the deceptions and noticed one or both of the relationships between the visual words and the loaded spoken words. They also correctly guessed that they were to respond in a certain way when conditioning was tested with the semantic differential. Most of these "aware" subjects responded by marking the critical words at the extreme ends of the seven-point scale. When asked to explain their responses, they typically said: "I thought that was what I was supposed to do; isn't that what you were trying to get me to do?" These subjects were called cooperators. Certain other subjects also verbalized awareness of the correct experimental contingency, but did not indicate it in their rating responses. In the post-experimental interview, these subjects said they tried to resist the association the experimenter was attempting to form. These subjects were called resisters. Many of the subjects (60 percent) did not reveal any awareness even on an extended post-experimental interview and this group showed no evidence of conditioning. In other words, post-experimental verbal reports of the subjects' awareness (i.e., ideas about the purpose of the experiment and what was expected of them) were highly correlated with behavior in this

situation. The distribution of responses was bimodal with the aware co-operators clustering around the most extreme conditioning score possible. The unaware subjects and aware resistors were distributed approximately normally around a conditioning score of zero. It was concluded that previous studies, because of poor interview techniques, had failed to detect all aware subjects. These few awares, responding at the extremes, were enough to raise the means in the predicted direction so as to produce an overall significant result. This effect had then been erroneously interpreted as evidence for conditioning without awareness.

The Staats and Staats experiment parallels the design of the Schafer and Murphy study quite closely. Both studies involved deception but did not include adequate procedures for checking on its effectiveness. Both studies involved a training series, where certain stimuli were contiguously paired with other stimuli which supposedly elicited pleasant or unpleasant affect in the subject and a deceptively introduced test for change in behavior as a result of the training. In both studies, the significance of the data is accounted for by a sub-group of subjects who respond at the extremes while some subjects did not show the effect at all. Results of both studies were interpreted as being similar to classical conditioning and thus nonvoluntary. Because of the similarities of the operations used in the two experiments, it is at least possible that similar psychological processes underly the behavior observed in both situations.

Investigators of learning and awareness have attempted, at an empirical level, to show that verbal conditioning occurs only in aware subjects who cooperate. Their findings suggest that what the subject in an experiment thinks the experimenter wants him to do is more

influential than what he is specifically told to do. While the learning and awareness literature has been concerned mainly with the issue of learning with or without awareness, it has implications that cut across a wide spectrum of current experimental methodology in psychology. At a more general level, other psychologists have been concerned with similar considerations to which we now turn.

Demand Characteristics

Within the context of a growing contemporary concern with the social psychology of the psychological experiment, Orne (1962) has advanced the concept of the "demand characteristics" of psychological experiments. Traditionally, psychologists have treated experimental subjects as if they were passive responders to stimuli. Orne finds this assumption difficult to justify. He asks: "How does the human subject perceive the laboratory situation? How does he define his role? What does he think he is supposed to do?" A person who agrees to serve as a subject in a psychological experiment implicitly agrees to perform a wide range of actions on request without inquiring as to their purpose or duration. The experimenter-subject relationship is very unique in that the experimenter has a high degree of control over the subject. The subject, possibly because of a high degree of respect for science or for more personal reasons such as "evaluation apprehension" (Rosenberg, 1965; Silverman, 1965), sees his role as that of being a good subject. In a sense, the subject's behavior in an experiment may be viewed as problem solving behavior. At some level he sees it as his task to discover what he is supposed to do (the experimenter's hypothesis) and to validate his role as a good subject by doing it.

Orne calls the totality of implicit or explicit cues which convey the experimental hypothesis to the subject, the demand characteristics of the experimental situation. These cues may be mediated by the experimenter's behavior or by the logic of the experimental procedure itself, such as a training-test sequence or a test-manipulation-retest sequence. Behavior in an experiment is viewed as determined by two sets of variables: (1) the experimental or independent variables and (2) demand characteristics as perceived by the subject in the experimental situation. Orne suggests that to the extent different experimenters or slightly different designs do not obtain the same results, they are mostly artifacts of demand characteristics. Demand characteristics artifacts may explain why so many experimental results do not generalize to behavior outside the laboratory.

The demands of an experiment are aspects of the situation such as instructions, treatment conditions, or behavior of the experimenter which might communicate implicitly or explicitly that a specific response is expected by the experimenter rather than alternative responses. Demand characteristics are the demands of the experiment as perceived by the subject. Perceived demand characteristics may be either correct (i.e., correlated with the experimenter's hypothesis) or incorrect. When they are correct, they are operationally equivalent to the concept of "awareness" as employed in the verbal conditioning literature.

It appears that the two formulations (awareness and demand characteristics) are actually describing the same psychological phenomena. The findings of awareness in verbal conditioning are merely specific experimental examples of the more general problem of demand characteristics in psychological experiments. It is not necessary to use two different languages to describe these phenomena.

Subsuming the concept of awareness under the broader concept of perceived demand characteristics allows the handling of all ideas or hypotheses that a subject might have with regard to the behavior expected of him in an experiment in one theoretical formulation. In this dissertation, the terms awareness and correct perceived demand characteristics are used interchangeably.

Subjects often have ideas about the behavior expected of them that are incorrect. It is also important to know about these. Incorrect perceived demand characteristics do not necessarily bias the data in the direction of the experimental hypothesis (i.e., increase the probability of a type one error) as do those which are correct, but they do contribute to a larger variance in experiments (i.e., an increase in the probability of type two error).

The demand characteristics formulation makes a good deal of sense out of the learning and awareness literature. Some subjects correctly perceive the demand characteristics of the situation; they become aware of the relationship between the experimenter's reinforcements and their own behavior and then demonstrate that they are good subjects by increasing their responses in the reinforced category.

In the language of the demand characteristics formulation, the significant results of a decade or more of verbal conditioning research may be attributable to demand characteristics variables (or artifacts) rather than experimental variables. If demand characteristics variables do account for the verbal conditioning effect, then this has broad implications for other areas of psychological experimentation. It raises an immediate question as to how many other studies, using different apparatus and apparently demonstrating different phenomena, might have had

designs with demand characteristics similar to those of verbal conditioning studies. If there are such studies, their data and conclusions are now suspect. It has already been indicated that the Schafer and Murphy (1943) experiment may be such a study.

CHAPTER II

PROBLEM

In Chapter 1, the possibility was explored that the classic Schafer-Murphy effect could be a result of subjects becoming aware or correctly perceiving the demand characteristics of the experimental situation and then cooperating by reporting more rewarded profiles. The general hypothesis was advanced that the Schafer-Murphy effect could be better accounted for by assessing the perceived demand characteristics than it can be by the original interpretation; that a genuine modification of perception has occurred as a result of the reward and punishment treatment. Previous studies have suggested that subjects may actually see both aspects of the reversible figures at the 1/3 second exposure speed (Rock & Fleck, 1950; Smith & Hochberg, 1954). If this is true, then demand characteristics, as formulated in Chapter 1 become a reasonable explanation for subjects decisions to report the rewarded names.

In previous studies of the Schafer-Murphy effect, all subjects have been given identical reward and punishment treatment, yet there is a wide variation in the tendency to report rewarded profiles. We have already indicated that individual differences in susceptibility to monetary reinforcement cannot account for this variance because of differences within subjects in the proportion of rewarded responses to the two ambiguous

test stimuli. Suppose that this variance were highly correlated with post-experimental reports, with subjects reporting that they thought the experimenter expected high rewarded response from them having high rewarded response, and subjects not having this idea having, as a group, a lower proportion of rewarded responses. If this were the case, then it would seem that the construct "perceived demand characteristics" better accounts for the data than does the construct "genuine perceptual re-organization as a function of reward and punishment."

Pilot Studies

In line with the above formulation, a pilot study was conducted. The study was done to familiarize the experimenter with any problems that might arise in attempting to obtain data on a larger scale, to aid in the development of the extended post-experimental interview for use in the experiment, and to see if the Schafer-Murphy effect could be replicated. Of course, the experimenter was also interested in assessing the validity of the demand characteristics formulation before launching a time consuming and expensive large sample factorial experiment.

The first pilot studies were conducted during the last week of the semester. While it may seem unimportant to state the time during the semester that the experiment was run, this led to the fortuitous discovery of a very important variable on which the Schafer-Murphy effect seems to depend. This will be described later.

The original Schafer-Murphy (1943) procedure was followed. On the first five pilot subjects, the effect was not obtained. It was then decided to try the Jackson (1954) procedure which left out the number guessing originally employed during training by Schafer and Murphy. For

the next six subjects, again no replication was obtained. The total number of previously rewarded test responses for the first 11 subjects was 58. The total number of punished test responses was 58. There was a total of 60 mistaken identifications. This is a relatively higher proportion of mistakes than was reported in previous studies.

At this point, it was discovered that a seemingly unimportant procedural detail employed in the earlier studies had been overlooked. Schafer and Murphy (1943) used improved "helping faces" outside of the tachistoscope for the subject to view between trials. Jackson (1954) does not say whether he used "helping faces," but he claimed to have replicated the earlier study carefully. He listed all of the points at which he changed the earlier procedure and he does not mention leaving out the "helping faces." Smith and Hochberg (1954) also used "helping faces."

It was decided to place helping faces on the wall of the experimental room, and to again attempt to obtain the effect. Six more subjects were run under the more precise replication, with striking results. The proportion of mistaken identifications dropped from 35 percent to 22 percent. This latter is comparable to the Jackson study (24 percent). The total number of previously rewarded responses was 55, while there were only 19 punished responses. This compares to the Schafer-Murphy data where a 54-13 split was obtained on five subjects, and to Jackson's data where there was a 50-23 split for six subjects in one reward order and a 55-16 split for six subjects in the other order.

The six subjects in the above replication were given an extensive informal post-experimental interview. These interviews were tape recorded without the subjects' knowledge. Four of the subjects seemed quite aware

of the purpose of the experiment. All of the above reported Schafer-Murphy effect (55-19 split) was accounted for by these four subjects. Both of the unaware subjects gave an equal number of rewarded and punished responses.

During the first week of the following semester, the main experimentation was begun. Fourteen subjects were run, but again there was no replication of the Schafer-Murphy effect. New and larger helping faces had been made which did not contain the extra eyes (see Figure 1, Chapter 1) outside of the actual profile. The helping faces were then changed back to those used in the first pilot study, but the effect was still not obtained on the next several subjects. At this time it was necessary to reformulate the problem in the light of these unexpected results.

Hypothesis One

Reflecting on the results of the pilot studies, it seemed that the differences observed between results obtained at the end of the previous semester and those obtained at the first of the next semester might be due to the differences in the subjects themselves. The former sample of subjects had completed an introductory course in psychology while the latter subjects had received no formal instruction in psychology. Schafer-Murphy and Jackson had used introductory psychology students, but they do not state whether the experiments were run early or late in the semester. Interestingly, the one study in the literature which did not replicate the original results (Rock & Fleck, 1950) used high school students, who more than likely had not completed a course in psychology, as subjects.

In pilot studies, the psychologically sophisticated subjects used such language as "the effects of reinforcement on perception," and "I confirmed the hypothesis," in describing the purpose of the experiment. On the other hand, most of the psychologically naive subjects were not able to describe the purpose of the experiment correctly. Most of them knew that two profiles had always won money and that two profiles were present in the ambiguous test situation, but they did not make any connection between the two contingencies. It seems that they simply did not have the necessary information available to perceive the experiment as an integrated whole. Two of these 14 subjects were aware (compared to four out of six among the sophisticated subjects). When they described their awareness, they used a different language, i.e., "The tester expected I would see winners better than losers."

The above considerations lead to the first major hypothesis of this dissertation: Psychologically sophisticated subjects will show the Schafer-Murphy effect to a significantly greater degree than will psychologically naive subjects. This prediction was made because it was expected that sophisticated subjects would be more likely to become aware of the purpose of the experiment. Also, awareness or correctly perceived demand characteristics was expected to lead to cooperation in the majority of subjects.

It could be that general sophistication factors such as age and number of years in college would also make for more awareness. Because of this possibility, it was considered necessary to build general sophistication into the experimental design. It was predicted that general sophistication factors would not be as important for obtaining the Schafer-Murphy effect as the specific factor of sophistication in the field of psychology.

Hypothesis Two

The second major hypothesis of this dissertation was formulated with regard to the general discussion of awareness and demand characteristics presented in Chapter 1. Specifically, the hypothesis was that: The tendency to report more rewarded profiles in the Schafer-Murphy situation will be highly correlated with post-experimental reports of awareness of the experimental hypothesis, and removal of the aware subjects from the data will result in a relatively symmetrical distribution of rewarded response scores around a mean of 50 percent.

A secondary prediction was that it would not be awareness per se that accounted for the Schafer-Murphy effect. It was predicted that it would be possible to distinguish a sub-population of aware-cooperating subjects from the protocols of post-experimental interviews, and that awareness plus cooperation would actually account for the Schafer-Murphy effect.

In addition to the two major hypotheses, several other side issues were proposed for study by means of questions on an extended post-experimental interview. These issues were as follows: (1) Do subjects learn that two profiles consistently win and two consistently lose in the Schafer-Murphy situation? (2) Is reported motivation for the money used in the training series related to the Schafer-Murphy effect? (3) Do subjects know that two profiles are being presented in the test series? (4) How often do subjects, who know of the reversible nature of the test stimuli, actually see both figures at 1/3 second exposure speed? (5) Do subjects report more winning profiles in the test situation because they expect that by such behavior they will continue to win money as in the training series?

CHAPTER III

METHOD

Selection Of Subjects

A total of 130 undergraduate students enrolled in psychology courses at the University of Oklahoma were used as subjects in the experiment. Ten of these were eliminated from the statistical analysis because they either misunderstood or ignored the instructions to report the name of one figure in the ambiguous test situation. Since the dependent measure was based upon subjects reporting only one name when presented with the ambiguous test figures, there was no basis for comparing the behavior of these ten subjects who called both names with that of the others. The characteristics of the ten subjects eliminated from the analysis will be presented separately in the chapter on results.

Of the 120 subjects actually used in the analysis, 24 were selected in a nonsystematic manner (from introductory psychology students) and assigned to a control group. The 96 subjects assigned to the experimental condition were drawn from four separate sub-populations within the larger population of undergraduate psychology students at the University of Oklahoma.

Group I (N=24) or the underclass-naïve sample, consisted of 14 freshmen and ten sophomores enrolled in their first course in psychology. These students had received a minimum of one and a maximum of five weeks

of instruction in psychology prior to their participation in the experiment. This group included 15 females and nine males. The median age of the group was 19 with a range from 18 to 21.

Group II (N=24) or the upperclass-naive sample, consisted of 17 juniors and seven seniors also enrolled in their first course in psychology. These students had received a minimum of one and a maximum of six weeks instruction in psychology prior to their participation in the experiment. This group included eight females and 16 males. The median age of the group was 20 with a range from 19 to 30.

Group III (N=24) or the underclass-sophisticated sample, consisted of 13 freshmen and 11 sophomores enrolled in a second course in psychology. All of these students had completed an introductory course in psychology within one year prior to their participation in the experiment and, in addition, had received a minimum of three and a maximum of six weeks instruction in a second psychology course. This group included 13 females and 11 males. The median age of the group was 19 with a range from 18 to 22. This group was more sophisticated than their underclass counterparts in introductory psychology in another, and perhaps very important sense; 71 percent of them had previously participated in at least one psychological experiment.

Group IV (N=24) or the upperclass-sophisticated sample, consisted of 13 juniors and 11 seniors. They were all upperclassmen with previous coursework in psychology, but they varied considerably with regard to the amount of previous coursework. The median number of previous semester hours in psychology was nine with a range from three to 18 hours. This group included seven females and 17 males. The median age of the group was 21 with a range from 19 to 31. This group was also sophisticated in

another sense in that 50 percent of them had participated in at least one psychological experiment.

Subjects from all groups except Group III knew in advance that they would receive two points added to their final course grade for participating in experiments. They could select the experiments they participated in or choose not to participate, however, they cannot be considered as volunteers. Group III did not expect to receive grade points for participation. An attempt was made to approximately equate Group III with the others with regard to extrinsic motivation to participate by offering them 90 cents in money for volunteering for the experiment. An attempt was made to check on the equality of the values of two points or 90 cents as motivation for participating in the experiment. All subjects in Group III and an equal number selected from the other groups were asked, after the experiment, which was more valuable to them. Of the subjects asked, 89 percent said that the 2 points had more value. Thus, if extrinsic motivation to participate in the experiment is important, Group III was not comparable to the other three groups.

In summary, it was not possible to match the four experimental sub-populations on all possible variables. However, the groups could be unambiguously classified in a 2 X 2 factorial design in terms of the variable of college classification (which, incidentally, could be labeled as "age" with very few exceptions) and that of psychological sophistication. The latter variable is the more important and for this main effect in the design, the matching is quite comparable. Table 1 presents a summary description of the subject samples. It also illustrates the factorial design used in the present experiment.

Table 1

Summary of Characteristics of the Subject Populations

		<u>Group</u>			
		Psychologically Naive	Psychologically Sophisticated	Totals	
Under-Classmen	<u>Group I</u>		<u>Group III</u>	<u>Underclassmen</u>	
	Freshmen	= 14	Freshmen	= 13	Freshmen = 27
	Sophomores	= 10	Sophomores	= 11	Sophomores = 21
	Females	= 15	Females	= 13	Females = 28
	Males	= 9	Males	= 11	Males = 20
	Md. Age	= 19	Md. Age	= 19	Md. Age = 19
	Age Range	= 18-21	Age Range	= 18-22	Age Range = 18-22
	Md.Hrs.Psy.	= 0	Md. Hrs.Psy.	= 3	
	TOTAL N	= 24	TOTAL N	= 24	TOTAL N = 48
Upper-Classmen	<u>Group II</u>		<u>Group IV</u>	<u>Upperclassmen</u>	
	Juniors	= 17	Juniors	= 13	Juniors = 30
	Seniors	= 7	Seniors	= 11	Seniors = 18
	Females	= 8	Females	= 7	Females = 15
	Males	= 16	Males	= 17	Males = 33
	Md. Age	= 20	Md. Age	= 21	Md. Age = 21
	Age Range	= 19-30	Age Range	= 19-31	Age Range = 19-31
	Md.Hrs.Psy.	= 0	Md. Hrs.Psy.	= 9	
	TOTAL N	= 24	TOTAL N	= 24	TOTAL N = 48
Totals	<u>Naive</u>		<u>Sophisticated</u>	<u>Experimental</u>	
	Underclass	= 24	Underclass	= 24	Underclass = 48
	Upperclass	= 24	Upperclass	= 24	Upperclass = 48
	Females	= 23	Females	= 20	Females = 43
	Males	= 25	Males	= 28	Males = 53
	Md. Age	= 20	Md. Age	= 20	Md. Age = 20
	Age Range	= 18-30	Age Range	= 18-31	Age Range = 18-31
	Md.Hrs.Psy.	= 0	Md.Hrs.Psy.	= 3	
	TOTAL N	= 48	TOTAL N	= 48	TOTAL N = 96

Control Group

<u>All Naive</u>	
Underclassmen	= 17
Upperclassmen	= 7
Females	= 11
Males	= 13
Md. Age	= 20
Age Range	= 18-26
TOTAL N	= 24

Apparatus And Materials

A standard Dodge tachistoscope was used for presentation of the stimuli. This instrument has a constantly lighted white background field located 36 inches from the viewing port. Stimuli were presented by means of mirrors at a viewing distance of 18 inches. The illumination of the stimulus field and the background field were equated for brightness. The instrument was calibrated for an exposure speed of $1/3$ second by means of a photo cell connected to an electronic timer.

The tachistoscope was placed on top of a large desk in an experimental room approximately 10 X 10 feet in size. This arrangement placed the viewing port of the tachistoscope at approximately eye level when the subject was seated at the desk. Large cardboard shields ($2\frac{1}{2}$ X 3 feet) were attached to either side of the tachistoscope to eliminate the possibility of the subject seeing the stimuli as they were being inserted or removed from the tachistoscope. The subject could, however, see the head and shoulders of the standing experimenter.

To eliminate any glare from the overhead lights in the room, a cardboard shield was constructed to extend out over the viewing port from the top of the tachistoscope. As an added precaution, the overhead light directly above the subject was left off during the experiment. Thus, while the subject experienced little or no glare from outside light sources, the overall illumination of the experimental room remained approximately that of an average illuminated room.

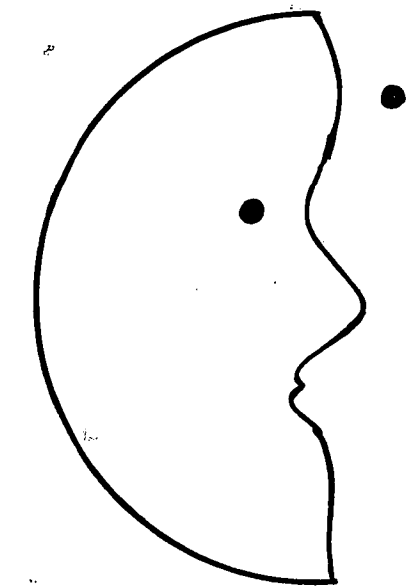
Eight stimulus figures, each having an outside diameter of three inches, were drawn in black ink on white cardboard cards $7\frac{1}{2}$ X 9 inches in size. These figures were reproduced as accurately as possible from those reported by Jackson (1954). On four of the stimulus cards were

drawn the half circle training figures illustrated in Figure 3. The names were not written under the figures for the actual stimuli. Two of the stimulus cards contained the ambiguous test figures illustrated in Figure 4. The remaining two cards contained the nonambiguous set breaking figures illustrated in Figure 5. In addition, four improved replicas or "help faces" similar to the training stimuli shown in Figure 2 were mounted on the wall of the experimental room at a viewing distance of approximately eight feet from the subject. Also, approximately eighty dollars in dimes was used over the course of the entire experiment as rewards and punishments for the subjects.

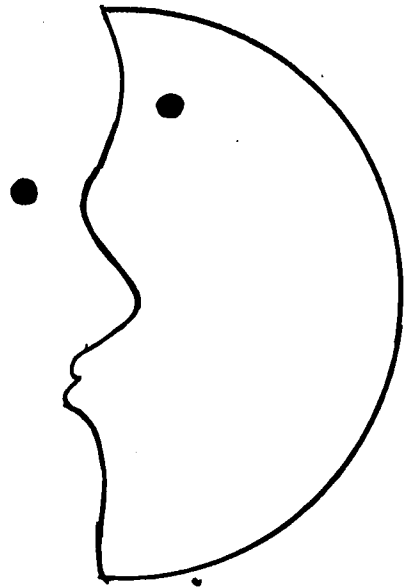
Experimental Procedure

It was difficult to determine from the Schafer-Murphy (1943) study the exact procedure at many points. A clearer statement of procedure, though different from the original at some points, was given by Jackson (1954). Since he achieved a successful replication, the changes in his procedure were considered as noncritical to the production of the Schafer-Murphy effect. The majority of later studies follow procedures similar to his. For these reasons, Jackson's procedure was followed as closely as possible for all four experimental groups in the present study, except that 10 cents instead of 15 cents was used as the reward. The use of a control group which learned the stimuli without the reward and punishment procedure necessitated a modification of instructions for that group which will be presented later.

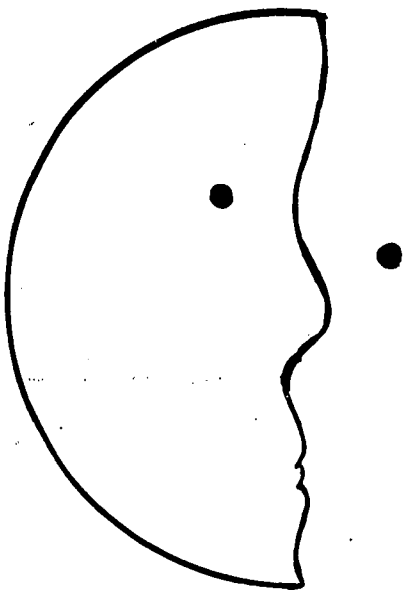
Experimental Groups. The subjects were ushered into the experimental room and told to be seated comfortably in front of the tachistoscope. Prior to their entry, two groups of dimes had been placed on the desk to



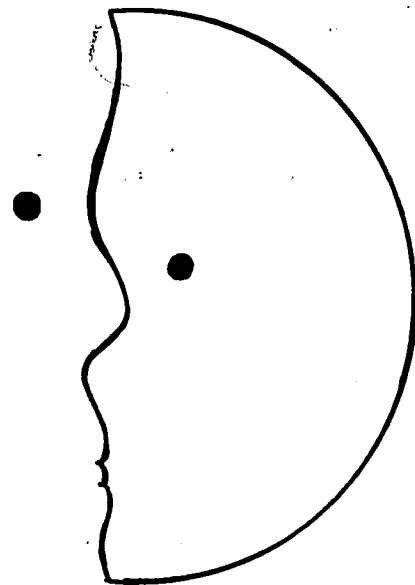
NATHAN



BERTRAND



CLIFFORD



DUNCAN

Figure 3. The four individual profile figures used in the training series.

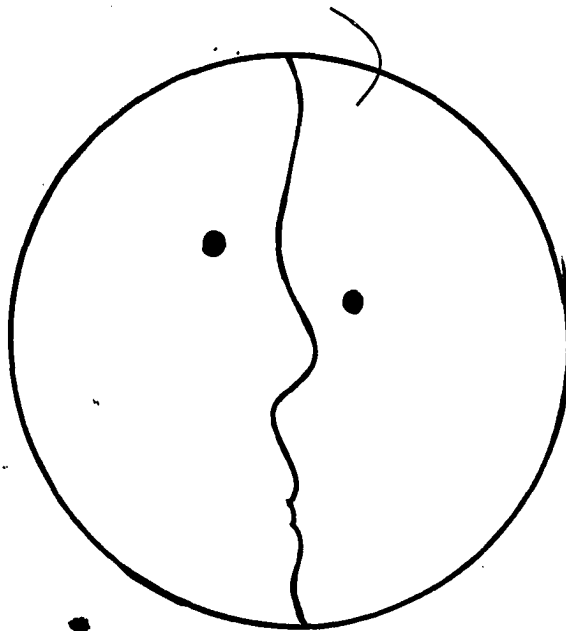
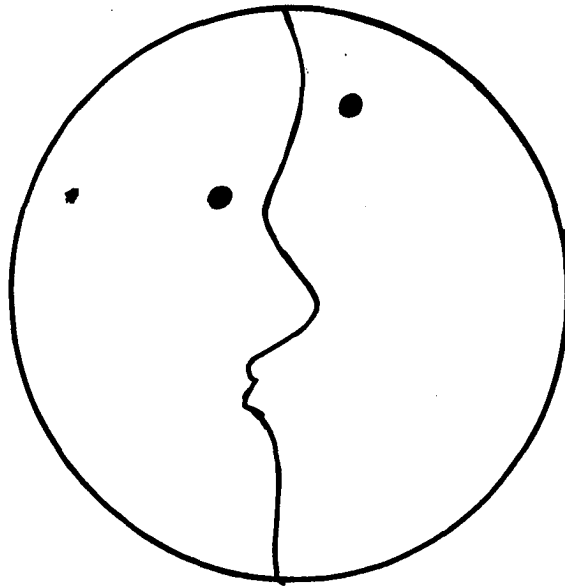


Figure 4. The two composite ambiguous profile figures used in the test series.

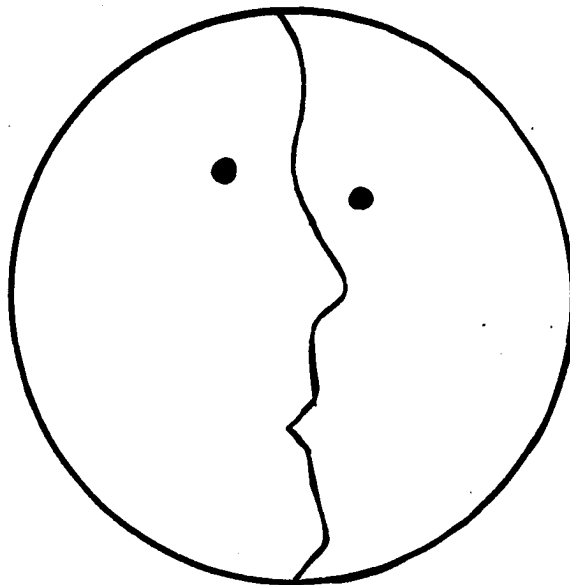
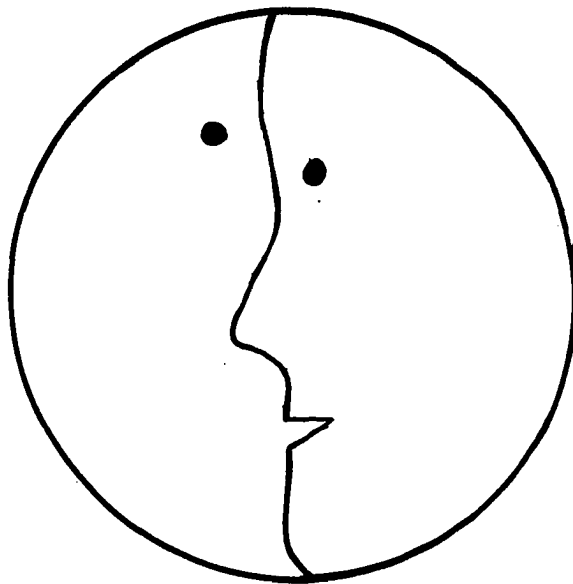


Figure 5. The two nonreversible set breaking profile figures used in the test series.

the subject's right. The group of dimes nearest the subject contained five dimes, the other group of dimes contained approximately 15 dimes.

The initial instructions for the experimental groups were as follows:

This is an experiment in vision. I am going to expose each of four profiles or faces rapidly a number of times in this machine. When either of two of the profiles is shown, you will actually win ten cents. When either of the other two is shown you will lose ten cents. I will start you off with fifty cents which you may consider as your pay for participating in the experiment. Your winnings and losings will be added to or subtracted from this amount. You may keep whatever you end up with. Also, each profile has a name and your task will be to learn the name that goes with each profile. To summarize what we shall do: (a) I shall announce the name of each face about to be presented. (b) You will observe the face in the machine. (c) You will repeat the name of the face aloud so that I will know that you have seen it and also this will help you to learn it. (d) I shall tell you whether you have won or lost ten cents. (e) You will take or return (demonstrating) ten cents and then we will go on to the next trial. Do you have any questions?

The training series was then begun. It consisted of 25 presentations of each of the four half-circle profile outlines (refer to Figure 2) totaling 100 training trials. After the first 50 trials and again at the end of 100 trials, a three minute rest period was given. This was shortened from the five minutes used in the original studies to three minutes when it was observed in the pilot studies that subjects seemed to become restless with the longer rest periods and typically attempted to engage the experimenter in conversation regarding the purpose of the experiment.

The rest periods seemed to serve two functions in the experiment. First, they gave both the experimenter and the subject a necessary rest from the tedious and prolonged test series. Second, they gave the subject an opportunity to study the "help faces" on the wall which, as suggested by the pilot work, both facilitated learning and gave the subject the time

needed to scan the profiles and thereby discover the possibility of the pieces fitting like a jigsaw puzzle. Both of these functions seemed to be adequately served by a three minute rest period.

During the training for a given subject, the winning of a dime (reward) was consistently paired with two of the profiles and the losing of a dime (punishment) was consistently paired with the other two profiles. The profiles receiving reward and those receiving punishment were systematically counterbalanced between subjects so that half of the subjects received rewards paired with Nathan and Duncan (N-D order) and the other half received rewards paired with Bertrand and Clifford (B-C order). This was done in an attempt to control for direction of facing and position of the eye dots for the different profiles (see Figure 3).

The sequence of presentations of the profiles was the same for all subjects during training. This sequence was determined by use of a table of random numbers, with two restrictions. First, the same profile could not appear twice in succession and second, there had to be an equal number of presentations of each of the four profiles within the series of 100 trials.

Because reward was consistently paired with two of the profiles and punishment with the other two, all subjects ended the training series with their original 50 cents. However, there was wide fluctuation in the amount of money that the subject had in his group of dimes over the course of training. The fluctuations ranged from 20 cents to one dollar.

During the first eight training trials, subjects attention was directed to the help faces on the wall before they viewed them in the tachistoscope. After the first eight trials, no further mention was made of the help faces. It was observed, however, that many subjects made frequent reference to these help faces between training trials, during

the rest periods, and before reporting their perception in the test series.

The tachistoscopic exposure speed for all phases of the experiment was 1/3 of a second. The experimenter made frequent checks between individual subjects to insure that all of the bulbs in the tachistoscope were working properly and that proper exposure speed was maintained.

Control Group. For the control group, 100 trials of training were also given, but the dimes were not present and no reference was made to the winning and losing of money. The initial instructions were as follows:

This is an experiment in vision. I am going to expose each of four profiles, or faces, rapidly a number of times in this machine. Each profile has a name and your task will be to learn the name that goes with each profile. To summarize what we shall do: (a) I shall announce the name of the face about to be presented. (b) You will observe the face in the machine. (c) You will repeat the name of the face aloud so that I will know that you have seen it and also this will help you to learn it. (d) Then we will go on to the next trial. Do you have any questions?

Test Trials. At the end of the second three minute rest period, identical instructions for beginning the test were given to subjects in all groups. For the experimental subjects, the money was left on the table and no further reference was made to the winning and losing of money. The following instructions were given:

Now we shall see how well you learned the faces and whether you can distinguish new faces from old. I am going to present the faces as I did before, except that this time they will be in a complete circle. All you have to do is to tell me the name of the face that you see. Quite frequently a new and strange face will be presented. When you do not recognize the face as one of those shown previously, you can indicate that the face is strange by saying "X." O. k.?

Originally it was planned to add the phrase: "Do you have any questions?" which was reportedly done by Jackson. However, in the pilot

work with naive subjects, this was found to be impossible. The typical question asked by these naive subjects was: "But there will be two faces if you put them in a circle. Am I supposed to name them both?" This put the experimenter in the position of having to deny that both would be present or of saying "No, just name one," or remaining silent. Therefore, it was decided to merely say "O. k." in a tone of voice that implied that no questions would be answered. This reduced the amount of questions asked but did not entirely eliminate the problem. Apparently, the instructions appeared ambiguous and misleading to several of the naive subjects and their impulse, quite justifiably, was to ask for clarification which often destroyed the intent of the deceptive instructions. When such incidents arose in the main experiment, the experimenter attempted to handle them by remaining silent when possible. If the subject insisted on a verbal answer, the experimenter attempted to be as noncommittal as possible. It is very interesting to note that this problem did not arise with the sophisticated subjects. They, typically, did not ask questions and had little apparent trouble in understanding and carrying out the instructions.

After the above instructions were given, a series of 30 test trials using the ambiguous or reversible test stimuli (refer to Figure 4) were begun. Between each test trial a nonreversible set breaking figure (refer to Figure 5) facing in the opposite direction from the profile the subject last reported seeing, was presented. Thus, there was a total of 60 presentations of stimuli in the test series.

During the test series, the experimenter remained seated behind the screen so that he could more easily record the subject's responses. In the training series, he had remained standing while operating the tachistoscope. Before each trial, the experimenter said, "ready" which

was the subject's signal to prepare to view the stimulus. After viewing the figure, the subject reported verbally the name of the figure he saw. These were recorded on a prepared data sheet by the experimenter. The order of presentation of the test trials was the same for all subjects. The order of presentation was determined from a table of random numbers with the restrictions that each of the two test stimuli occurred five times in each block of ten test trials and a single test figure could not appear more than twice in succession.

Extended Interviews. After the test series, all 96 experimental subjects were given an extended post-experimental interview questionnaire. The 24 control subjects were given an abbreviated form of the same questionnaire. Based on the reticence of pilot subjects to engage in an oral interview, it was decided to give the interview in written form. This form also allowed for control over variations in the questioning procedure. The questions were presented individually on half sheets of paper stapled together in booklet form. Subjects went through the booklet answering one question at a time. They were not allowed to look ahead to the next question until they had completed the earlier one and they were not allowed to go back and add anything to previous answers once they had turned the page.

The questionnaire was carefully constructed so that clues as to the nature of the experiment were not given until the later part of the questionnaire. It was intended that a subject's correct verbalization of the purpose of the experiment (his awareness) expressed at the first of the interview would be weighted heavier in scoring the interview than would the same verbalization expressed at the end of the questionnaire. Instructions for the extended interview questionnaire were as follows:

We now come to the most important part of the experiment. We want to find out what you thought the experiment was about while it was going on. I have here a booklet containing several questions concerning what you thought about during the experiment. Please answer each question in order, before going on to the next question. Once you have turned a page, do not go back and change your answers to previous questions. Please answer each question as conscientiously and truthfully as possible, because, as I said, this is the most important part of the experiment. Keep in mind that we want to know about things you thought of during the experiment and not things that might occur to you in retrospect as you are writing. Feel free to explain your answers in detail. Later you will have a chance to ask questions and to find out what this is all about.

The questionnaires for both the experimental and the control groups may be found in Appendix A.

The door to the experimental room was opened so that the subject could not refer to the helping faces on the wall. After the subject had completed the questionnaire, he was given his fifty cents in winnings (or in the case of Group III, ninety cents) and allowed to ask any questions that he might have regarding the experiment. The experimenter revealed all deceptions and discussed the overall purpose of the research candidly. The importance of not discussing the experiment with anyone for the duration of the semester was stressed. Each subject then signed a written statement to the effect that he would not discuss the experiment with anyone.

The length of time required for the entire experimental session varied from 55 to 65 minutes depending on how much time the subject spent in filling out the questionnaire.

CHAPTER IV

RESULTS

Scoring Procedures

As a first step in the analysis of the data, it was necessary to decide on an adequate method of scoring the test responses. Schafer and Murphy (1943), Jackson (1954), and others who have worked with this phenomena considered each individual test response for each subject as an independent event. They then summed (across subjects) the total number of correct test responses which had been previously rewarded and also the total number of correct responses which had been previously punished. Incorrect responses (the calling of Clifford, Duncan, or "X" when the Nathan-Bertrand test stimuli was presented, etc.) were left out of the analysis. Schafer and Murphy then applied a one sample proportion test to these data while Jackson used a chi square goodness of fit test. Both of these statistical procedures are illegitimate because the frequencies are not based upon independent observations. Each subject contributed a number of responses to the total of both the rewarded and punished categories. The proportions of rewarded versus punished responses between subjects were grossly unequal and within subjects, they were dependent; i.e., as reward goes up, punished must go down.

Thus, the problem in scoring the data in the present experiment was to obtain an index of tendency to report rewarded profiles which

would be comparable to the measures used previously and yet would be more adequate for statistical analysis. It was deemed desirable to have one score for each subject, representing the tendency to report rewarded profiles, only for those trials on which he gave a correct response. As in previous studies, the number of mistakes for any given subject was considered primarily a function of individual visual acuity.

The required score was arrived at for each subject by dividing the total number of times that the names of the two rewarded profiles were correctly reported during the test series by the total number of correct responses. That score, in turn, was multiplied by 100 to obtain a percentage of rewarded responses score. These scores were then used in the analysis. All subjects performed at a level better than 50 percent correct responses. The range of mistakes was from zero to 15 of the 30 trials and the median number of mistakes was three.

Another problem in scoring concerned the number of test trials over which the behavior was to be scored. Schafer and Murphy based their analysis only on the first 16 test trials because they found the effect only up to that point. Jackson did not find a change in results when all of the test trials were scored. A preliminary look at the present data, which is based on five times the number of subjects used in both previous studies combined, did not reveal any substantial differences between using the first 16 trials as an index or using all of the test trials. It can be argued that a percentage score based on a possible 30 responses is a more stable index than one based on only 16, and therefore the analysis presented in this chapter is based on the former index. An explanation for Schafer and Murphy's loss of statistical significance after the first 16 trials can probably be found in examining their use of

a small sample of subjects in terms of awareness. This problem will be taken up in the discussion chapter.

In order to compare the control group to the other groups, it was necessary to arbitrarily assign half of the control subjects to one reward order (N-D) and the other half to the other order (B-C). The control subjects were then scored in the same manner as the experimental subjects.

Next, the post-experimental interviews for awareness were carefully read and scored for awareness-unawareness by four independent judges. One of these judges was the experimenter and the other three were advanced graduate students in psychology at the University of Oklahoma. These latter three judges had no knowledge of the subjects behavior on the test. Judges were instructed in detail as to the procedures of the experimental session. They were told to weigh the first part of the interview more heavily in their judgments than the later part. Judgments were made, forced choice, for aware-unaware. Further instructions were given to the effect that only subjects who were clearly aware of the experimenter's hypothesis were to be classified as aware. Subjects who's awareness was doubtful were scored as unaware. As a further precaution against attributing awareness to a subject where such an attribution might be questionable, only those 39 subjects on which there was 100 percent agreement among the four judges were considered as aware for purposes of data analysis.

The judges also rated each subject on a five-point scale for cooperation-noncooperation. This data was for use in the correlational and nonparametric analyses.

Hypothesis One

Recall the first hypothesis advanced in Chapter 2 that psychologically sophisticated subjects would show the Schafer-Murphy effect to a

greater degree. Relevant data are presented in Table 2 which displays the mean percent of rewarded responses for the four experimental groups and the control group. Table 3 presents the corresponding analysis of variance. It can be seen that the naive subject population ($\bar{X} = 52.13$) differs significantly ($F = 4.59, P < .05$) from the sophisticated population ($\bar{X} = 60.25$), supporting the first hypothesis.

In addition to the factors shown in Table 2, the data were also partitioned in terms of reward order (N-D versus B-C order). The contrasts below the line in Table 3 are not orthogonal to the sources of variance above the line. These later contrasts were pulled out on a-priori grounds in order to pinpoint some of the sources of variation in the experiment.

Table 3 also shows that there is no significant difference between underclassmen ($\bar{X} = 55.31$) and upperclassmen ($\bar{X} = 57.65$). This variable was inserted into the design as a control variable. From this nonsignificant result we can infer that it is more than likely not general sophistication factors such as number of years in college and age that make for the increased Schafer Murphy effect. This in turn, lends credence to the possibility that the effect is mediated by a more specific factor such as knowledge of the field of psychology. Because of the lack of significance, the college classification factor will not be included in further analysis of the data.

The contrast between the control group and all experimental groups was not significant. This is not surprising since the mean of the two naive groups ($\bar{X} = 52.13$) is quite close to that of the control group ($\bar{X} = 50.75$). When the naive ($\bar{X} = 52.13$) and the sophisticated ($\bar{X} = 60.25$) groups are contrasted separately with the control, the difference is

Table 2

Mean Percentage of Rewarded Responses for Five Groups in the Experiment

	Psychologically Naive	Psychologically Sophisticated	Totals
Upper- Classmen	50.54	60.08	55.31
Lower- Classmen	53.71	60.42	57.65
Totals	52.13	60.25	56.19
Control Group			
	50.75		

Table 3

Analysis of Variance for the Data in Table 2

<u>Source of Variance</u>	MS	df	F	P
Naive-Sophisticated	1584.37	1	4.59	< .05
Underclass-Upperclass	73.49	1	-	NS
Interaction	48.16	1	-	NS
Control-Experimental	567.68	1	1.64	NS
Order for Control	53.50	1	-	NS
Error	345.25	110	-	
<u>Non-Orthogonal Contrasts</u>				
N-D Order - B-C Order	9511.80	1	27.55	< .001
Sophisticated-Control	1444.00	1	4.18	< .05
Naive-Control	30.25	1	-	NS

significant ($F = 4.18$, $P < .05$) for the sophisticated group but not for the naive. Thus, a replication of the Schafer-Murphy effect was obtained only with psychologically sophisticated subjects.

The difference between reward orders comes out as the most significant source of variation ($F = 27.55$, $P < .001$) in the experiment. This result was entirely unanticipated. This effect does not appear in the control group where the scoring of reward order was arbitrary. This suggests that something other than figure dominance is occurring in the experimental groups.

Table 4 presents the means of the five groups of the experiment when the data is broken by reward order. It may be seen that three of the four experimental groups have N-D order means that are lower than the control mean for N-D order. All of the B-C means are above the A-D and the control means. This puzzling order effect will be further analyzed later in this chapter.

Hypothesis Two

A second major hypothesis proposed in Chapter 2 was that there would be a highly significant positive correlation between subjects' verbalization of awareness of the experimental hypothesis and tendency to report rewarded profiles during the test.

The judges agreed 100 percent as to aware-unaware for 77 of the 96 experimental subjects. The 19 subjects on which agreement was less than 100 percent were about evenly split with regard to those scored aware by three judges and unaware by the fourth and those scored unaware by three judges and aware by the fourth. However, they were composed of more sophisticated subjects ($N=14$) than naive subjects ($N=5$). Of the 77

Table 4

Mean Percentage of Rewarded Responses
for the Five Groups in the Experiment,
Breaking by Reward Order

	Psychologically Naive		Psychologically Sophisticated		Totals
Under- Classmen	N-D*	41.46	N-D	47.30	55.31
	B-C	65.67	B-C	69.21	
Upper- Classmen	N-D	42.33	N-D	56.08	57.65
	B-C	65.83	B-C	65.55	
Totals		52.13		60.08	56.19

Control Group

N-D 49.10
B-C 52.25

Total 50.75

*Reward Order

subjects unanimously classified as either aware or unaware, 39 were aware and 38 were unaware. The point biserial correlation between awareness and behavior on the test was $r_{pb} = .52$ which was highly significant ($t = 5.29$, $P < .001$, $df = 75$).

Another way of illustrating the same relationship is to classify the data in a 2 X 2 contingency table according to the variables of aware versus unaware and percent rewarded score of 50 or below versus those above 50 percent. Such a classification is presented in Table 5. Of the 39 aware subjects, only eight (21 percent) scored below 50 percent, while 23 (61 percent) of the unaware subjects scored below 50 percent. The chi square test of this association was highly significant ($X^2 = 12.81$, $df = 1$, $P < .001$).

Thus, the second major hypothesis, that awareness is related to tendency to give high percentages of rewarded responses in the Schafer-Murphy situation, is strongly supported by the data.

Recall from Chapter 2 the proposal that it is not the awareness per se which mediates the higher percentage of responses in the direction of the experimenter's hypothesis among aware subjects. Rather, when awareness is present, the subject has the alternative of either cooperating or not cooperating. It was proposed that the decision to cooperate with the experimenter is the important factor in producing the behavior. If this proposal were correct, then the strength of the awareness effect should be improved by singling out only aware cooperating subjects and comparing them to all others.

The judges classified 29 of the 39 aware subjects as at least partial cooperators. The other ten awares were classified as either not cooperating or deliberately resisting. When the 29 cooperating subjects

Table 5

Frequencies of Rewarded Response Scores Above and Below
50 Percent for Aware and Unaware Subjects

	Aware	Unaware	Totals
Above	31	15	46
Below	8	23	31
Totals	39	38	77

$$(\chi^2 = 12.81, df = 1, P < .001)$$

versus all others were correlated with behavior, the correlation increased to $r_{pb} = .70$ which is highly significant ($t = 8.49$, $df = 75$, $P < .0001$). This represents a substantial increase in prediction over the correlation based only on awareness alone ($r_{pb} = .52$).

Figure 6 presents the frequency distribution of all 48 unaware and noncooperating subjects separately from that of the 29 cooperating aware subjects. The medians of these two distributions are quite different (48 versus 81). It may be seen that removing the aware cooperators from the data completely eliminates the upward skew in the overall distribution. The distribution of aware cooperators is very similar to the overall distributions reported in the Schafer-Murphy (1943) and the Jackson (1954) studies.

The increase in predictability that occurs when cooperation-noncooperation is taken into account may be further illustrated by classifying only the 39 aware subjects on this variable and comparing them on behavior (see Table 6). The chi square test for this association is highly significant ($\chi^2 = 12.86$, $df = 1$, $P < .001$).

The two cooperating subjects who were below 50 percent rewarded responses were special cases. Each of them reported that they did not become aware until about the middle of the test series, at which time they began cooperating. Inspection of their responses shows that they responded predominantly with punished profile names for more than the first half of the test trials. They then suddenly reversed themselves and reported nothing but rewarded profile names for the remainder of the test. One of these subjects exclaimed, "Oh!" after the 17th trial. He never missed reporting a rewarded name after that trial.

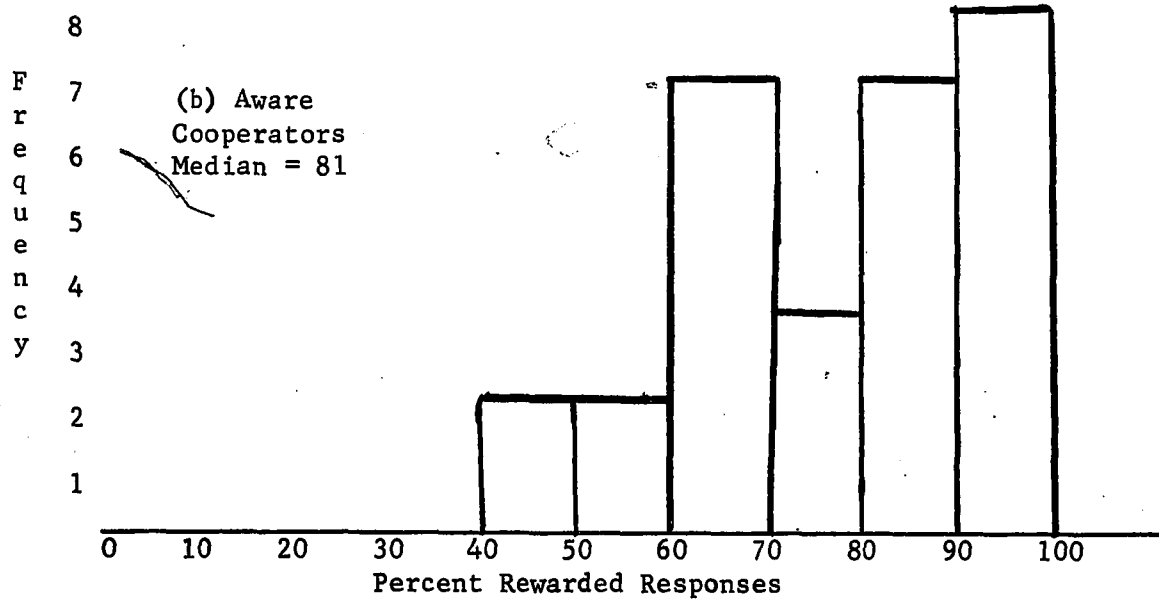
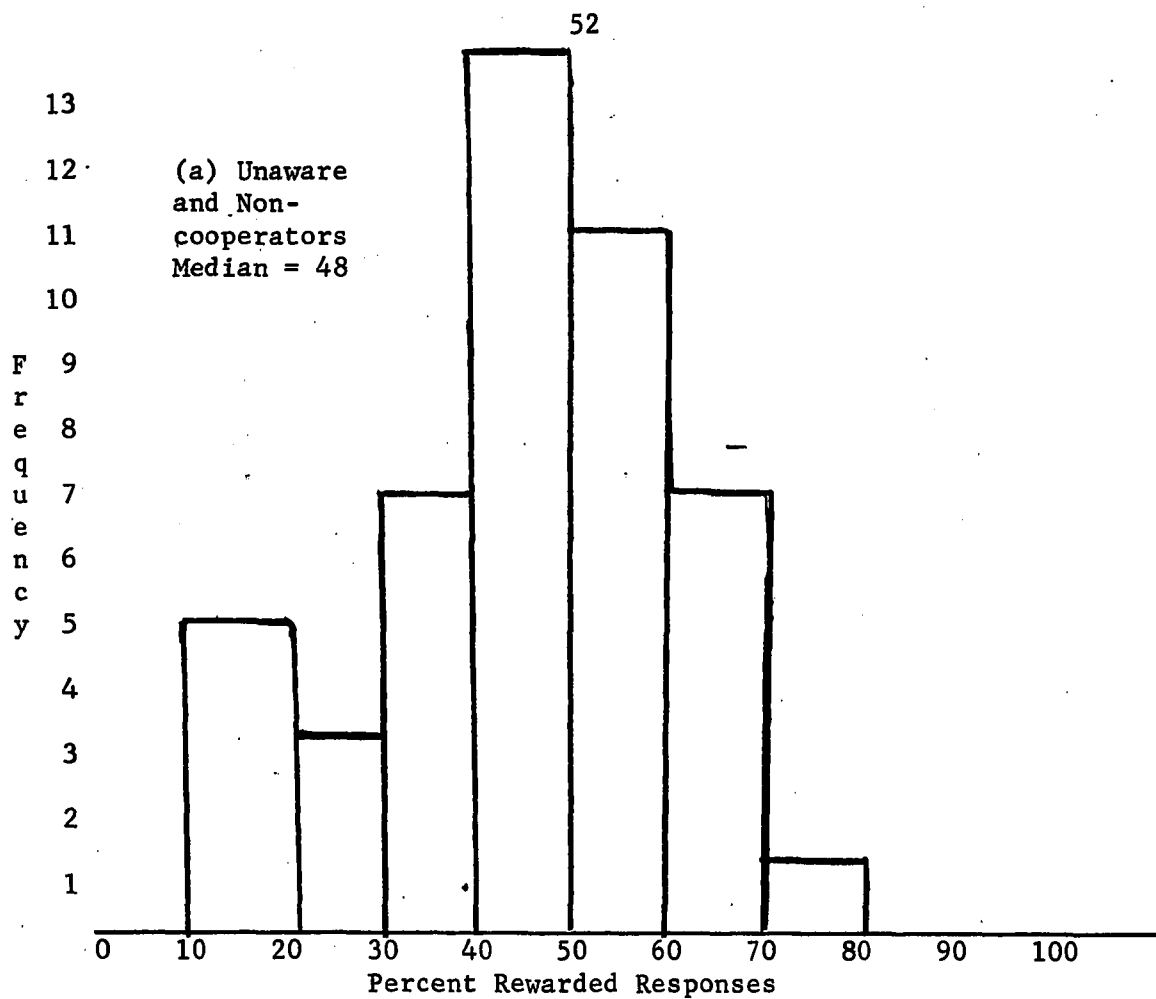


Figure 6. Frequency distributions for (a) the 48 unaware and noncooperating subjects and (b) the 29 aware cooperating subjects.

Table 6

Frequencies of Aware Cooperators and Noncooperators
Who Score Above and Below 50 Percent
Rewarded Responses

	Cooperate	Noncooperate	Totals
Above	27	4	31
Below	2	6	8
Totals	29	10	39

$$(\chi^2 = 12.86, df = 1, P < .001)$$

Internal Analyses

The Order Effect. The significance of the order effect found in the main analysis, i.e., subjects in the B-C order scored higher, necessitated further analysis of the data. Because it was suspected that awareness had something to do with this effect, only those 77 subjects on which inter-judge agreement was obtained on awareness were used in this analysis. The control group which did not show a significant order effect and could not be classified with regard to aware-unaware was also not included in this analysis.

In order to evaluate the possibility that an unequal proportion of recognitions for a single figure was producing the order effect, the data was rescored in terms of percentage of correct responses of each rewarded figure separately.

Table 7 shows the mean rewarded and punished responses obtained under reward, control, and punished conditions for each of the faces. The control means are included for comparison purposes although they were not included in the analysis. In this table, the same subjects contribute a score to each of four of the cells. For example, the scores in the cells Nathan-Duncan rewarded and Bertrand-Clifford punished are contributed by the same subjects. Also, the responses were made to only two ambiguous test situations (Nathan-Bertrand and Clifford-Duncan). Therefore, the diagonals for these two blocks necessarily add to 100 percent. As a result, only the scores in the rewarded part of this table were included in the analysis of variance shown in Table 8. Inspection of the table reveals that it is only the means of the scores of the 38 subjects in the Bertrand-Clifford reward order that do not compare closely to the respective control group means. Yet, their scores compare closely among

Table 7

Mean Percentage of Rewarded and Punished Test Responses
Scored Separately for Each of the Four Figures

	Figures				
	<u>Nathan</u>	<u>Bertrand</u>	<u>Clifford</u>	<u>Duncan</u>	Totals
Rewarded	50.41 N=39	64.89 N=38	66.63 N=38	44.69 N=39	56.54 N=154
Control	47.87 N=24	52.13 N=24	53.38 N=24	46.62 N=24	50.00 N=96
Punished	35.11 N=38	49.59 N=39	55.31 N=39	33.37 N=38	43.46 N=154

Table 8

Summary of the Analysis of Variance for Awareness, Subject Population,
Reward Order, and Figures Within Order for the 77 Subjects
Who Could be Unambiguously Classified on Awareness*

Source of Variance	Adjusted Mean Square	df	F	P
Awareness	5169.6	1	5.36	< .05
Sophistication of Population	788.8	1	.82	NS
Reward Order	6404.3	1	6.64	< .05
Awareness x Population	1186.4	1	1.23	NS
Awareness x Order	257.7	1	-	NS
Population x Order	37.8	1	-	NS
3-way Interaction	91.6	1	-	NS
Figures Within Order	143.8	2	-	NS
Figures x Population	2261.9	2	1.75	NS
Figures x Awareness	1219.0	2	-	NS
3-way Interaction	1280.6	2	-	NS
Error	964.5	69	-	
Nested Error	1226.6	69	-	

*Method of unweighted means (Winer, 1962).

themselves. This is an important point that will be taken up later in the nonparametric analysis.

In Table 8 it may be seen that sophistication of population which was the significant variable in the first analysis (Table 3, page 48) does not approach significance in this analysis. This is as it should be if support for the awareness hypothesis is to be obtained. Recall that the reason for predicting a significant sophistication of population effect was that there would be more awareness in the sophisticated group. If this were the case, then when awareness and sophistication were pulled out as independent factors, the awareness factor should be significant (and it is, $F = 5.36$, $P < .05$) while the subject population factor should not be significant ($F = .82$, NS). The interaction between awareness and subject population ($F = 1.23$, NS) also is not significant. Therefore, support is also found for the awareness hypothesis in this analysis. This analysis makes clear the source of the significant difference found between subject populations in the experiment, i.e., the sophisticated group scores higher because more of them are aware.

Again, this analysis indicates a significant reward order effect ($F = 6.64$, $P < .05$). The significance of this effect is somewhat less pronounced in the second analysis. Apparently some of this effect was absorbed in the higher-order interactions (none of which reached significance). The figures within order effect does not approach significance. Thus, the possibility that the order effect is the result of a single figure being dominant is ruled out. The second analysis yields some new information, but the order effect remains unexplained.

A further exploration of the order effect was carried out by the use of nonparametric techniques. When the 77 individuals utilized in the

above analysis are classified according to the variables reward order and scores above and below 50 percent rewarded responses (see Table 9), a highly significant association ($\chi^2 = 22.91$, $df = 1$, $P < .001$) is found. This is essentially the same relation as the order effect in the analysis of variance.

Further examination of the data at least partially accounts for the magnitude of the above result. Recall from Table 5 (page 50) that awares score high while unawares score low. Table 10 shows that there is a greater proportion of awares in the B-C group (61 percent) than in the A-D group (41 percent). This relation approaches statistical significance ($\chi^2 = 2.93$, $df = 1$, $P < .10$), and could account for some of the order effect.

Cooperation among aware subjects has already been found to be more predictive of high rewarded responses than awareness alone (see Figure 5, page 35). If the B-C order which includes more aware subjects also included a higher proportion of cooperators, then this would further account for the order effect. Table 11 in fact, shows that there is a significantly ($\chi^2 = 4.67$, $df = 1$, $P < .05$) greater proportion (87 percent) of cooperators in the B-C order than in the N-D order (56 percent).

The data in Table 7 now makes sense. The scores for the Bertrand and Clifford means were contributed by the same subjects. More of these subjects were aware and more of them cooperated. Therefore, the means are high. The scores for the Nathan and Duncan means were contributed by another group of subjects. In this group there were fewer awares, and of these awares, seven did not cooperate. Of these seven noncooperators in the N-D group, six scored below 50 percent rewarded responses. The scores of four of these were individually significant (binomial test, Walker & Lev, 1953) in the opposite direction from the hypothesis (30

Table 9

Frequencies of Subjects Scoring Above and Below 50 Percent
Rewarded Responses in the Two Reward Orders

	Reward Order		
	<u>N-D</u>	<u>B-C</u>	Totals
Above 50 percent	13	33	46
50 percent or below	26	5	31
Totals	39	38	77

$(\chi^2 = 22.91, df = 1, P < .001)$

Table 10

Frequencies of Aware and Unaware Subjects in the Two Reward Orders

	Reward Order		
	<u>N-D</u>	<u>B-C</u>	Totals
Aware	16	23	39
Unaware	23	15	38
Totals	39	38	77

$(\chi^2 = 2.93, df = 1, P < .10)$

Table 11

Frequencies of Cooperating and Noncooperating Aware Subjects
in the Two Reward Orders

	Reward Order		Totals
	<u>N-D</u>	<u>B-C</u>	
Cooperate	9	20	29
Noncooperate	7	3	10
Totals	16	23	39

$(X^2 = 4.67, df = 1, P < .05)$

percent or below). The net result of the factor of awareness in the N-D group was that the cooperators and the resisters balanced each other, and the mean percent rewarded scores remained similar to those of the control group.

In summary, it seems that the order effect is associated with unequal distributions of awares and unawares and cooperators and noncooperators in the two reward orders. Whether or not the B-C order actually facilitated awareness is impossible to determine from the data. It seems more likely that it was the unequal proportions of awares in the two groups which produced the order effect rather than the order effect producing the awareness because the order effect is not significant in the control group.

Other Data. Some other factors relevant to the Schafer-Murphy effect were also examined in the extended interview of the present experiment. The original interpretation of the effect was that the pleasant experience associated with winning money was mediating the behavior. According to this interpretation, subjects who score high should be the same subjects who were motivated by the money. Question seven of the extended interview (Appendix A) was intended to distinguish between subjects who were motivated by the money and those who were not. A point biserial correlation was computed between this measure and percentage of rewarded responses on the test. The correlation coefficient for this test was $r_{pb} = .136$ which is not significantly different from zero. If we assume that asking subjects a yes-no question concerning their interest in winning the money is even a rough index of degree of motivation for the money, then there appears to be no relationship between desire for money and the Schafer-Murphy effect. In contrast to this correlation, recall that percentage of rewarded responses was highly correlated ($r_{pb} = .70$,

$P < .001$) with awareness plus cooperation. The latter is an index of motivation to please the experimenter and to be a good experimental subject.

The original Schafer and Murphy experiment was conducted under the assumption that subjects would not be able to perceive both aspects of the ambiguous reversible figures at 1/3 second exposure speed. In the present experiment, subjects were asked in the extended interview, concerning their knowledge that the figures fit together and also how often they actually did see both profiles together on the test. Of the total of 130 subjects run in the experiment, only 14 could not match up the correct faces which fit together when asked to do so in the extended interview. Six of the 14 were in the control group. All of the remaining 116 (or 89.2 percent) subjects answered "yes" to the question: "Did you ever notice or suspect that there were two profiles inside the circle on the test, depending upon which way you looked at it?" and then went on to correctly match up the profiles. Some of these subjects only suspected that there might be two profiles together, but never really saw two together clearly. Table 12 shows the frequency distribution of subjects ratings on a five-point scale concerning their certainty that two profiles were present. Notice that 59.2 percent of them were absolutely certain that two profiles fit together on the test.

Table 12 also shows the data for the question: "The faces which always won money during learning were?" and the frequencies of certainty ratings for this question. Control subjects did not answer this question so the total N is 106. Notice that 82.9 percent did learn which profiles always won money and could correctly recall them after an interval of 60 test trials on which no winning and losing took place. Most (72.6 percent of the total N) were entirely certain of this.

Table 12

Frequency Distribution of Subjects Ratings of Certainty
That Two Profiles Fit on the Test
and Which Two Profiles Won Money

Certainty of Figures Fitting						Correct Total	Incorrect Total
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>		
Frequency	6	7	5	21	77	116	14
Percent	4.6	5.4	3.9	16.1	59.2	89.2	10.8
						Total N = 130	
Certainty of Winning Figures							
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>		
Frequency	0	1	3	7	77	88	18
Percent	0	.9	2.8	6.6	72.6	82.9	17.1
						Total N = 106	

When aware and unaware subjects are compared on relative certainty that the figures fit on the test (Table 13), there is a significantly greater proportion ($X^2 = 8.78$, $df = 1$, $P < .01$) of awares who were certain of this contingency. Also, comparing awares and unawares on certainty of the figures which won money (Table 14) shows that there is a significantly ($X^2 = 15.02$, $df = 1$, $P < .001$) higher proportion of awares who were completely certain of the winners. It appears that subjects who knew which profiles won and were relatively certain that a winner and a loser appeared together on the test were more likely to become aware.

The above analysis does not take into account the frequency of subjects who actually saw two profiles together on the test. A subject could know or suspect that two profiles were being presented, yet one of them could be dominant at the 1/3 second exposure speed. The subject would then be reporting only one name because he actually only saw one profile. The extended interview contained a question which helps evaluate this possibility. Subjects were asked to circle one of five alternatives with regard to the number of times they actually saw two profiles together on the test. The alternatives were: (1) never, (2) a few times, (3) several times, (4) often, and (5) always. The most reasonable place to partition these ratings into a saw few-saw many dichotomy seems to be between ratings two and three. Table 15 shows that a significantly greater proportion of awares ($X^2 = 9.7$, $df = 1$, $P < .01$) claimed to have seen two profiles together on the test. It appears reasonable to assume that subjects who often saw two profiles had to make a decision as to which name to report. It may be that having to make this decision facilitated awareness.

Table 13

Frequency of 39 Aware and 38 Unaware Subjects Classified
According to Relative Certainty or Uncertainty
That Figures Fit

	Aware	Unaware	Totals
<u>Certain</u>			
4-5 Ratings	35	23	58
<u>Uncertain</u>			
1-3 Ratings and Mistaken Identification	4	15	19
Totals	39	38	77

$$(X^2 = 8.76, df = 1, P < .01)$$

Table 14

Frequency of 39 Aware and 38 Unaware Subjects Classified According
to Relative Certainty or Uncertainty of the Two Profiles
Which Always Won Money

	Aware	Unaware	Totals
<u>Certain</u> 5 Ratings Only	35	19	54
<u>Uncertain</u> 1-4 Ratings and Mistaken Identifications	4	19	23
Totals	39	38	77

$$(X^2 = 15.02, df = 1, P < .001)$$

Table 15

Frequency of 39 Aware and 38 Unaware Subjects Classified According
to Whether They Actually Saw Two Profiles Together on the Test
a Few or Many Times

	Aware	Unaware	Totals
Saw Few	9	22	31
Saw Many	30	16	46
Totals	39	38	77

$$(X^2 = 9.7, df = 1, P < .01)$$

Data was collected with regard to another factor which could possibly be related to the Schafer-Murphy effect. Subjects may have reported more winning profiles on the test because they expected that they would win more money by doing so. On the extended interview, subjects were asked if they had at first expected to continue winning money on the test. Table 16 shows that there was no association between behavior on the test and expectancy of winning more money ($\chi^2 = .009$, $df = 1$, NS).

It could be that awareness is related to the expectancy of receiving more reward on the test. However, Table 17 reveals no significant association between awareness and this expectancy ($\chi^2 = 1.28$, $df = 1$, NS). In Table 18 it is interesting to note that there is a significant relation ($\chi^2 = 4.57$, $df = 1$, $P < .05$) between the expectancy of more money and cooperation-noncooperation when only awares are considered. By inspection of the table, we see that this is due to nine out of ten of the noncooperators not expecting more money on the test. This could be a clue to the noncooperating group having a negative attitude toward the experiment, although there is no way of investigating this possibility.

Subjects Who Reported Both. Recall from Chapter 3 that ten subjects were eliminated from the main analysis because they consistently called the names of both profiles in the ambiguous test situation. All of these subjects reported that they always saw both profiles together on the test when asked about this on the extended interview. However, 67 other subjects also reported being certain of seeing both profiles on the test yet they responded as instructed with the name of only one. Previous studies do not report difficulty with subjects reporting both names on the test. For this reason, data from these ten deviant subjects were examined more closely. All ten of these subjects were unaware. Under the

Table 16

Frequency of 96 Subjects Who Scored Above and Below 50 Percent
Rewarded Responses Classified According to Whether or not
They Expected to Win More Money on the Test

	Expectancy		Totals
	Yes	No	
Above 50 Percent	26	30	56
50 Percent or Below	18	22	40
Totals	44	52	96

($\chi^2 = .009$, $df = 1$, NS)

Table 17

Frequency of 39 Aware and 38 Unaware Subjects Classified According
to Expectancy of Winning More Money on the Test

	Expectancy		Totals
	Yes	No	
Aware	15	24	39
Unaware	10	28	38
Totals	25	52	77

($\chi^2 = 1.28$, $df = 1$, NS)

Table 18

Frequency of 29 Cooperating and 10 Noncooperating Awares Classified
According to Expectancy of Winning More Money on the Test

	Expectancy		Totals
	Yes	No	
Cooperate	14	15	29
Noncooperate	1	9	10
Totals	15	24	39

$$(x^2 = 4.57, df = 1, P < .05)$$

hypothesis that the population probability of a subject being aware is .50, which is reasonable from the present data ($A=39$, $UA=38$), the binomial probability of all ten of these subjects being unaware is .001 (Walker & Lev, 1953). Therefore, it seems unlikely that "both-saying" subjects were unaware by chance. Also, nine of the ten subjects were from the naive subject population. Under the hypothesis that the probability of a "both-saying" subject being drawn from either population is .50, the binomial probability of a nine-one split is .022. Accordingly, it seems unlikely that these subjects came from the naive group by chance.

Among the 96 experimental subjects retained in the analysis, only seven match the description of the nine naive subjects described above, that is, having always seen both, being unaware and naive. Therefore, out of this specific sub-population, "both-saying" was not uncommon, in fact it occurred more than 50 percent of the time. For the other 89 subjects, it occurred only once. If previous experimenters did not draw a sample from the population of naive unawares with superior visual acuity, then it is not surprising that they did not encounter the "both-saying" problem.

CHAPTER V

DISCUSSION

The Present Findings

Recall from Chapter 2, the general objectives of this dissertation: first, to test the hypothesis that the Schafer-Murphy effect would be found to a greater degree using a psychologically sophisticated group of subjects than with a comparable sample of psychologically naive subjects; second, to test the hypothesis that the Schafer-Murphy effect could be accounted for by a subgroup of subjects who were aware of the experimenter's hypothesis and who were cooperating in their responses. Strong support for both of these predictions was obtained in the experimental results. The data are clear: subjects who think that the experimenter expects them to pick previously rewarded profiles comply with these expectancies in the majority of cases. Also, the probability that a subject will correctly perceive the experimenter's purposes is significantly increased if he has at least an elementary knowledge of the field of psychology.

Subjects who were not aware of the experimenter's hypothesis varied considerably in their behavior on the test. Some consistently reported the names of both profiles. Others developed directional sets, that is, consistently naming the profile on the right or left. Some others reported alternating with one name and then the other. Those unawares with less than average visual acuity had no problem in deciding which name

to call as they apparently saw only one profile. This variability balanced out over a large sample of unaware subjects so that there was no more tendency to report previously rewarded profiles than there was to report the punished ones.

Some subjects were difficult to classify with regard to awareness. They seemed to know quite a bit about the purpose of the experiment but they thought that the experimenter could be interested in discovering which profiles, either the rewarded or the punished, would be dominant in perception. These subjects apparently were not operating under the demand to report rewarded profiles. Among these subjects, some went on to report that while they thought the experimenter could expect either one, he probably expected the rewarded profiles to be dominant. Most of these latter subjects were among the 19 on which the judges could not agree as to whether they were aware or not. The former subjects, who reported only that either one could have been expected, were classified as unaware. Other subjects ($N = 3$) reported that they thought the experimenter expected punished profiles to be emphasized in perception. Interestingly, all of these had percent rewarded response scores below 50 percent.

The finding of a significant difference between psychologically naive and sophisticated samples is particularly interesting. In fact, the naive sample in the present study, while containing some aware co-operating subjects, did not differ significantly from the control group. Not only did these groups differ in the behavior measured in the experiment, but they differed in other ways. Nine of them did not conform to the instructions to report only one profile on the test. Many of them asked questions such as: "What is this experiment about?" and seemed amused

or embarrassed by the reward and punishment procedure. The sophisticated subjects, on the other hand, typically did not ask questions and followed instructions without hesitation.

The naive aware subjects did not have the psychological language for describing their awareness. They frequently referred to the experiment as a "test," and used such terms as "Bertrand was a good guy while Nathan was the bad guy," or "Good ole Bertrand won me money," etc. The sophisticated aware subjects gave qualitatively different post-experimental reports. They used words like: "hypothesis," "reinforced stimulus," "reward and punishment," "conditioning," "perception," and "I chose the profiles on which I had received reinforcement." Obviously, subjects who have had a course in psychology do learn something. Yet, 90 percent of the data in psychology is collected on subjects taking a course in psychology.

Use Of Post-Experimental Interviews

The present research was conducted under the assumption that at the present level of development of psychology, even a rough index of something psychologically meaningful can have value while something precisely operationalized and formulated in mathematical language can be based on a behavior that is an artifact of an experimental situation which has no parallel outside the experimental setting and therefore may be worthless (Heider, 1958).

It was further assumed that the best method of finding out what subjects think about during a controlled experimental session is the most obvious way, that is, to ask them afterwards in a way that lets them understand that you really want to know and that you expect them to be honest. Some psychologists have hesitated to accept such an approach for fear that

there is too much slippage between what subjects are able to verbalize at a later point in time and the actual cognitions they might have had. There is, admittedly, some slippage in the method of post-experimental reports, but this does not invalidate their scientific value if they are used with reasonable and objective controls. Their scientific value is increased when they are put to the test of inter-scientist agreement.

As an illustration of the kinds of post-experimental reports given by various kinds of subjects after the experiment, selections from some actual protocols are presented below.

In the classical verbal conditioning of meaning (Page, 1964), the dependent measure was taken only once, consequently cooperation was mostly all or none. In the present study, the dependent measure was taken a total of 30 times, and some subjects claimed to cooperate for awhile and then to pick losers for awhile. The following quote from a male subject in Group IV illustrates the reports of a partial cooperator.

(Question one. Purpose of the experiment?) To see if reinforcement effects perceptual learning. The reinforcement in the preliminary learning session appeared, at least to me, to prompt almost automatically, the reinforced response when the actual test began. (Question three. Purpose of putting the profiles inside a complete circle.) To see which face I, the subject, would pick out, i.e., perceive. (Question four. On the test the experimenter expected) the faces of Bertrand and Clifford, the reinforced responses. These responses were reinforced. Therefore, it can only be expected that these would turn up. One would also expect the "punished" responses to turn up less often. (Question 18. Make any additional statements which apply to your behavior on the test.) I tried to cooperate for awhile; then I decided to be a little mischevious by picking out a losing face now and then.

This subject was a junior and had completed nine undergraduate hours in psychology. He was aware and cooperating, but not always. His percent rewarded response score was 67.

Some aware subjects claimed to have cooperated almost completely. This was not confined to the sophisticated group, though there were more of them in that group. To illustrate these points, the following quote is from a male sophomore who had not previously taken a course in psychology. He reported 100 percent rewarded profiles on the test.

(Question one. Purpose of the experiment.) To try to see whether or not learning can be induced by rewarding a person for his responsiveness. Also, to see whether or not the person would see the faces that meant money to him. (Question three. Purpose of putting the profiles inside a complete circle.) The two faces in the circle were one which would lose money and one which would win. The subject would be tested to see if he identified the pictures that would win for him. (Question four. On the test, the experimenter expected) faces that made me money. This would determine whether or not earning money was a stimulus to learning. (Question 18. Make any additional statements which apply to your behavior on the test.) I knew, because of the training period, that Bertrand and Nathan and Clifford and Duncan went together because of the lip-like appearances on the profiles. Therefore, I answered with the winning profiles.

To complete the illustrations of subjects answers to the extended interview questions, the protocol below was given by the male junior who had no previous hours in psychology and was classified as unaware.

(Question one. Purpose of the experiment.) During the first part of the experiment I had little idea of what I was doing. The second part showed me that my mind could not react instantly to all of the things it saw, and neither could it remember them all. I guess I thought the purpose of the experiment was to see how much the mind could register and how much it could remember to answer correctly. (Question three. Purpose of putting the profiles inside a complete circle.) This, I believe, was to see if when the profile was within a complete circle the subject would tend to try to see two profiles or just see the ones he had been seeing during the experiment. (Question four. On the test, the experimenter expected) I believe that the experimenter wanted me to see one of the profiles I had seen during the experiment and to pick out the ones I had not yet seen. The experimenter probably expected that I would see many new profiles as the profiles with the complete circle around them created many new profiles. (Notice that he doesn't seem to know that there are two of the previously learned profiles within the circle. This subject was one of the 14 who could not correctly match

up the profiles which fit on question 12) (Question 18. Make any additional statements with regard to your behavior on the test.) I wasn't really sure of how to react at times. On such a test, it is easy to miss, as all these profiles start to look alike. It is a very hard test, I believe, but I enjoyed doing it.

While it wasn't possible to obtain precise interval scale measurement of the subjects' awarenesses or response hypotheses, study of their reports makes it abundantly clear that what the subjects thought about during the experiment were powerful determinants of behavior. Even the rough classifications of aware-unaware and cooperate-not cooperate accounted for all of the systematic variance in the experiment.

Interpretation Of Previous Literature

It now appears reasonable to suggest that past reports of modifying perception by the use of small amounts of money contiguously with training on certain of the stimuli were probably produced by aware and cooperating subjects. It seems that past experimenters were modifying verbal reports in their testing situation, but the underlying psychological processes involved may well have been quite unrelated to these experimenter's interpretations of the behavior. Small amounts of money given and taken away do not seem to modify what is seen when an ambiguous figure is perceived. What is reported as seen is, however, modified under the special circumstance where the subject thinks that the experimenter expects that he will be able to see those figures better for which he has received money.

In the present experiment, it was observed that unaware subjects were not always consistent in their strategies for choosing one profile over the other throughout the test. Some aware subjects tired of cooperating consistently over 30 trials, while others did not become fully

aware until toward the end of the test at which time they began cooperating. Recall that Schafer and Murphy obtained the effect only until about the 16th trial at which time some of their subjects reversed their bias from rewarded to punished profiles. Their other subjects continued to report rewarded profiles. If we assume that some of their subjects were aware while others were unaware (or that all were aware), then their loss of the effect after 16 trials is not surprising. It could be that their awares continued to cooperate, while their unawares just happened to be going predominantly with the rewarded profiles at first. We cannot attribute this effect to the rewards, because of the small sample ($N = 5$). With a larger sample, they may have gotten some unawares who would have reported the punished profiles. Jackson (1954) did have three subjects in 12 who did not show the effect and three more who gave only a slight preference to rewarded profiles. If these were unaware, then they did, in fact, balance out. Jackson's other six subjects gave between 75 and 100 percent rewarded responses and resemble closely the aware cooperators in the present study.

Concerning the differences between psychologically naive and sophisticated subjects, it is interesting that the only other recorded instance of negative results with the Schafer-Murphy effect occurred when a sample of high school students served as subjects (Rock & Fleck, 1950). It is true that a projector tachistoscope was used and that subjects probably could not see the stimuli as well. However, Smith and Hochberg (1954) also used a projector tachistoscope with significant results. Jackson (1954) plays down his results obtained with a projector tachistoscope in an effort to explain away Rock and Fleck's negative results on the grounds that they used the wrong apparatus. However, Jackson did obtain 59 percent

rewarded responses summed over 18 subjects. He did get 70 percent rewarded responses using a Dodge tachistoscope. But, even this might not mean that the apparatus produced the difference. In the light of the present findings, this difference could be accounted for if Jackson ran his subjects with the projector earlier in the semester.

Schafer and Murphy (1943) report that they did not get the effect in a pilot study, but they got tendencies to develop directional sets. They then invented their set breaking figures. The five subjects run later, using set breaking figures, did show the effect. In the present study, the set breaking figures did not appear to work very well in that many unaware subjects developed directional sets. Rock and Fleck (1950) formally analyzed the effect of the set breaking figures and found them to be ineffective. One can speculate as to the events that might have transpired in psychology class between Schafer and Murphy's pilot study and the study they report. It could well be that a chapter on learning theory was studied. Actually, Schafer and Murphy report that two of their subjects were not in a psychology course. These subjects are not identified. One of them could have been their control subject. It would be interesting if these were the subjects who changed from rewarded to punished responses after 16 trials on the test, or were the two subjects with the lowest number of rewarded responses.

Solley and Murphy (1960) discuss the literature on reinforcement and figure-ground perception in detail. Concerning awareness, they say that the effect is best obtained at low levels of awareness. Specifically, they say:

Subjects are aware of the perceptual materials and of the reward and punishment stimuli, but they are rarely aware of the connection between the two. -- it appears that maximal autistic

perception is obtained if the reinforcement or punishment stimuli are barely in awareness.

The present results directly contradict this claim. In the present study it was found that the effect is maximally obtained at very high levels of awareness. It seems strange that Solley can make any statement about awareness, because nowhere does he report the results of a formal study of awareness. It may be that Solley's subjects did not report high levels of awareness simply because he never really asked them what they were aware of.

While past experimenters have asked a few questions concerning subjects' awarenesses in the Schafer-Murphy situation, what was asked (or how) is not usually reported. Jackson (1954) asked two questions. First, "What was the purpose of the experiment?" and second, "Were you interested in winning the money?" Concerning his first question, it has been found, in verbal conditioning studies, that such a question does not detect more than half of the aware subjects. Jackson does report one subject as saying, "The purpose was probably to get an association of winning 15 cents with Nathan and Duncan." This was one of Jackson's four extreme responders. It is likely that an extended interview similar to that used in the present study would have revealed that the other three extreme responders were just as aware. Concerning Jackson's other question, he found, "Since it was clear that rewards and punishments were having some effect upon perception," that further "probing" was necessary. By this, we assume that subjects who had initially said that the money didn't make any difference were asked again, "Oh! Come now, you really did want that money, didn't you?" It is not surprising that aware cooperating subjects, believing that

the experimenter must expect that they should have been interested in the money, would then acquiesce to such strong social pressure and say, "Well, maybe I did want it a little." This is a clear cut example of a demand characteristic.

The criticism that demand characteristics also influenced the results of the present post-experimental interview could also be raised. However, there are some important differences. First, a specific attempt was made in the present experiment to release subjects from any demand characteristics by the instructions before the interview. They were told, in effect, that the experimental game that the experimenter had previously required them to play was now over and that the task was now to honestly report what they had thought about during the game. Second, while the interview was extended, it was not at all "probing" in the sense that Jackson uses that term. In short, the demands were to tell the truth and not what the experimenter might expect to hear. Third, the interview was written which eliminated the possibility of the experimenter conveying his expectancies through tone of voice or facial expression. In the present situation it would seem that a subject still operating under the demands of the "game" would see it as his task not to reveal his awareness rather than to "ruin the experiment" by confessing that he knew what it was about.

Implications For The Social Psychology Of Experiments

Psychologists who used experimental procedures such as those employed in studies of reward and punishment on perception must have been either unaware of or ignoring the large number of possible social influence variables present in such a situation. While some of these

variables cannot be precisely measured, they cannot be ignored. Rosenthal's (1963) experimenter bias, Spielberger's (1962) awareness, and the present study's concern with cooperators suggest that these are factors that may be mediating behavior in many experiments.

Communication does not require specific explicit verbalization. Aside from the logic implicit in the situation itself, the experimenter may communicate much information concerning his expectancies by tone of voice and gestures (Rosenthal & Fode, 1963). The present experimenter has never conducted an experiment where some subjects didn't ask questions which required deviation from the planned procedure, yet these are never discussed in formal reports of results. The use of deception, when it is not effective, may also induce in the subject the set to begin trying to figure out the experiment. All of these things can communicate to an experimental subject the experimenter's expectancies. In the Schafer-Murphy situation, the experimenters would have saved time and money had they merely told their subjects that they were to "Pick Nathan instead of Bertrand when you see them together." The same instruction is apparently communicated by having the subject push money back and forth.

The experimenter is an authority figure, a "psychologist," and usually older than the subject. Because of this, the subject, who knows (or thinks he knows) what the experimenter wants of him, may feel strong social pressure to conform to his wishes. In any experiment of the type described in this dissertation, it seems legitimate to ask whether the behavior is a function of the experimenter's carefully measured independent variables or of the experimenter's wishes. The only alternatives are to either isolate and control for all of these factors or to find other methods of studying behavior that are not subject to these problems.

The finding that there were differences between psychologically sophisticated and psychologically naive subjects also has implications that go beyond the specific experimental setting in which it occurred. It seems that subjects who believed they were supposed to be conditioned behaved as if they were. Subjects who hadn't studied psychology were less likely to have acquired these beliefs. It is a common practice among psychologists to use undergraduate psychology students as subjects. In the light of the present findings, it is suggested that this practice should be re-evaluated. Psychologists should consider the scientific generality of data supporting their theories, when that data has been obtained from people already trained to believe that these theories are true.

CHAPTER VI

SUMMARY

In recent years, two related problem areas have arisen. The first has been concerned with the problem of awareness in verbal conditioning experiments. The second concerns the problem of demand characteristics in psychological experiments and more generally the social psychology of the psychological experiment. The present study points out that these two problems are actually concerned with the same psychological phenomena. The former is merely a special experimental example of the latter more general problem.

The specific research problem was concerned with the well-known Schafer-Murphy effect of reward and punishment on figure-ground perception. It was proposed that the experimental effect was an artifact of demand characteristics, or of subjects becoming aware of the experimenter's hypothesis. This possibility was investigated using 96 experimental subjects who responded to extended post-experimental interviews concerned with awareness of the purpose of the experiment. Pilot research suggested the hypothesis that psychological sophistication of subjects might be related to a tendency to report rewarded profiles, because of greater occurrence of aware subjects in this population.

Strong support was obtained for both hypotheses. Many subjects reported that they actually saw both of the profiles in the reversible

figure-ground situation and had to make a decision as to which name to report. Of these, more than half reported that they thought the experimenter expected them to report rewarded names. The Schafer-Murphy effect was accounted for by a subgroup of highly aware subjects who respond consistently with the names of previously rewarded profiles in the ambiguous figure-ground test. Unaware subjects showed no more tendency to report rewarded than punished profile names. There was a significant difference between a group of psychologically naive subjects ($N = 48$) who were just beginning a course in introductory psychology and a group of psychologically sophisticated subjects ($N = 48$) who had completed such a course the previous semester or the previous year. Only in the latter group was a replication obtained. General sophistication factors such as college classification and age were not significant, suggesting that the difference between groups was due to the specific factor of knowledge of the field of psychology.

The finding of a relationship between awareness of the purpose of the experiment and the Schafer-Murphy effect was discussed in terms of social psychological factors typically ignored in laboratory experiments. In line with the finding of differences between psychologically naive and sophisticated subjects, it was suggested that psychologists re-examine the common practice of teaching their theories to students and then testing these theories using the same people as subjects.



REFERENCES

- Ayllon, T. and Sommer, R. Autism, emphasis, and figure-ground perception. Journal of Psychology, 1956, 41, 163-176.
- Dollard, J. and Miller, N. E. Personality and psychotherapy. New York: McGraw-Hill, 1950.
- Dulany, D. E. Hypotheses and habits in verbal "operant conditioning." Journal of Abnormal and Social Psychology, 1961, 63, 251-263.
- Dulany, D. E. The place of hypotheses and intentions: an analysis of verbal control in verbal conditioning. In C. W. Eriksen (Ed.) Behavior and awareness. Durham, N.C.: Duke University Press, 1962.
- Farber, I. E. The things people say to themselves. American Psychologist, 1963, 18, 185-197.
- Jackson, D. N. A further examination of the role of autism in a visual figure-ground relationship. Journal of Psychology, 1954, 38, 339-357.
- Greenspoon, J. The reinforcing effect of two spoken sounds on the frequency of two responses. American Journal of Psychology, 1955, 68, 409-416.
- Heider, F. The psychology of interpersonal relations. New York: John Wiley & Sons, Inc., 1958.
- Letchworth, G. A. The classical conditioning of meaning as a function of awareness and need for social approval. Master's Thesis, University of Oklahoma, 1965.
- Levin, S. M. The effects of awareness on verbal conditioning. Journal of Experimental Psychology, 1961, 61, 67-75.
- Levine, R., Chein, E., and Murphy, G. The relation of the intensity of a need to the amount of perceptual distortion: a preliminary report. Journal of Psychology, 1942, 13, 283-293.
- Murphy, G. and Solley, C. M. Learning to perceive as we wish to perceive. Bulletin Menninger Clinic, 1957, 21, 225-237.

- Orne, M. T. On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. American Psychologist, 1962, 17, 776-783.
- Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. The measurement of meaning. Urbana: University of Illinois Press, 1957.
- Page, M. M. The role of awareness in the classical conditioning of meaning. Master's Thesis, University of Oklahoma, 1964.
- Proschansky, H. and Murphy, G. The effects of reward and punishment on perception. Journal of Psychology, 1942, 13, 295-305.
- Rock, I., and Fleck, F. S. A re-examination of the effect of monetary reward and punishment in figure-ground perception. Journal of Experimental Psychology, 1950, 40, 766-776.
- Rosenberg, M. J. When dissonance fails: on eliminating evaluation apprehension from attitude measurement. Journal of Personality and Social Psychology, 1965, 1, 28-42.
- Rosenthal, R. On the social psychology of the psychological experiment: the experimenter's hypothesis as unintended determinant of experimental results. American Scientist, Vol. 51, No. 2, 1963, 268-283.
- Rosenthal, R. and Fode, K. L. Psychology of the scientist, V. three experiments in experimenter bias. Psychological Reports, 1963, 12, 491-511.
- Santos, J. F. and Garvin, E. A. A further examination of the Schafer-Murphy effect. American Journal of Psychology, 1962, 75, 259-264.
- Santos, J. F. and Murphy, G. An odessey in perceptual learning. Bulletin Menninger Clinic, 1960, 24, 6-17.
- Schafer, R. and Murphy, G. The role of autism in a visual figure-ground relationship. Journal of Experimental Psychology, 1943, 32, 335-343.
- Silverman, I. Motives underlying the behavior of the subject in the psychological experiment. Paper presented at the A. P. A. Convention, 1965.
- Snyder, F. W. and Snyder, C. W. The effects of monetary reward and punishment on auditory perception. Journal of Psychology, 1956, 41, 177-184.
- Smith, D. E. P. and Hochberg, J. E. The effect of "punishment" (electric shock) on figure-ground perception. Journal of Psychology, 1954, 38, 83-87.

- Solley, C. M. and Engel, Mary. Perceptual autism in children: the effects of reward, punishment, and neutral conditions upon perceptual learning. Journal of Genetic Psychology, 1960, 97, 77-91.
- Solley, C. M., and Murphy, G. Development of the perceptual world. New York: Basic Books, Inc., 1960.
- Solley, C. M. and Santos, J. F. Perceptual learning with partial verbal reinforcement. Perceptual and Motor Skills, 1958, 8, 183-193.
- Solley, C. M. and Sommer, R. Perceptual autism in children. Journal of Genetic Psychology, 1957, 56, 3-11
- Sommer, R. and Ayllon, T. Perception and monetary reinforcement: I. the effects of rewards in the tactual modality. Journal of Psychology, 1956, 42, 137-141.
- Spielberger, C. D. The role of awareness in verbal conditioning. In, C. W. Eriksen (ed.) Behavior and awareness. Durham, N. C.: Duke University Press, 1962.
- Staats, A. W. and Staats, C. K. Attitudes established by classical conditioning. Journal of Abnormal and Social Psychology, 1958, 57, 37-40.
- Staats, C. K. and Staats, A. W. Meaning established by classical conditioning. Journal of Experimental Psychology, 1957, 54, 74-80.
- Taffel, C. Anxiety and the conditioning of verbal behavior.. Journal of Abnormal and Social Psychology, 1955, 51, 496-501.
- Tatz, S. J. Symbolic activity in learning without awareness. American Journal of Psychology, 1960, 73, 239-247.
- Walker, H. M. and Lev, J. Statistical inference. New York: Holt, Rinehart and Winston, 1953.
- Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill Book Co., 1962

APPENDIX A

EXTENDED INTERVIEWS

(Experimental Group)

- 89

(12) If so, draw lines connecting the profiles that fit together. If uncertain, guess.

Nathan	Nathan
Bertrand	Bertrand
Clifford	Clifford
Duncan	Duncan

(13) If you ever saw two profiles together on the test, what was the basis of your answer? In other words, why did you report one or the other names instead of reporting that you saw both? Explain:

(14) How often did you see two profiles together on the test?

Circle one: never, a few times, several times, often, always

(15) In this experiment, two profiles always won and two always lost in the training phase. In the test phase, two profiles always appeared together in the circles. One was a winning profile and the other a losing profile. It was expected that you would more often see the winning profile. Did something like this occur to you during the test?

Circle one: yes, no. How certain are you of your answer to number 15? (Mark X on one line.) uncertain _____ certain

(16) At first, on the test, did you think that you might win more money if you identified the winning faces or something similar?

Circle one: yes, no. Explain: In other words, what did you think about at the first of the test?

(17) Mark an X by any of the following statements that applies approximately to you. You may mark as many as you wish.

___ I had no idea of what the experiment was about during the test.

___ I thought I knew what the experiment was about, but my idea was not similar to that stated in number 15.

___ I knew approximately what the experiment was about on the test and when there was any doubt as to which face I saw in the circle, I reported a winning face.

___ I knew approximately what the experiment was about on the test, but I called them the way I saw them. If I saw both, I either said both or X. Which? _____

___ I knew what the experiment was about, but I tried to resist being influenced. If anything, I reported more losing than winning faces on the test.

___ When there were two faces in the circle, I always gave the name of the one on the right or left (circle one) and did not pay any attention to the winning or losing part of it.

___ The winning faces really did stand out more clearly than the losing ones.

___ It was hard to decide which stood out.

___ I sometimes noticed both profiles, but I didn't know if I was supposed to report both.

___ I deliberately tried to cooperate, in that while knowing that two faces were present, I reported the winning face because I thought that was what I was supposed to do.

___ I think I saw the winning faces more often because I was looking for them.

___ I noticed that Nathan and Bertrand could be seen together but I didn't notice Clifford and Duncan.

___ I noticed that Clifford and Duncan could be seen together but I didn't notice Nathan and Bertrand.

(18) Make any additional statements, not included above, which apply to your behavior on the test.

(Control Group)

(1) The purpose of this experiment was:

(2) The purpose of putting the profiles inside a complete circle on the test was:

(3) On the test, the experimenter expected that I would see:
Please explain your answer to question 3.

(4) Did you ever notice or suspect that there were two profiles inside the circle on the test, depending upon which way you looked at it?
Circle one: yes, no.

(5) If so, how certain were you of this? (Mark an X on one line.)
uncertain _____ certain

(6) If so, draw lines connecting the profiles that fit together. If uncertain, guess.

Nathan	Nathan
Bertrand	Bertrand
Clifford	Clifford
Duncan	Duncan

(7) If you ever saw two profiles together on the test, what was the basis of your answer? In other words, why did you report one or the other names instead of reporting that you saw both? Explain:

(8) How often did you see two profiles together on the test?
Circle one: never, a few times, several times, often, always