

SELF-PERCEPTION THEORY AND CREDIBILITY
CUEING: CONCEPTUAL AND EMPIRICAL ANALYSES

by

Ronald Lew Douglas

A DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

In the Department
of
Psychology

We accept this ~~dissertation~~ as ~~conforming~~ to the
required ~~standard~~

THE UNIVERSITY OF BRITISH COLUMBIA
March, 1974

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

Department of Psychology

The University of British Columbia
Vancouver 8, Canada

Date May 22/74

ABSTRACT

The theories of cognitive dissonance (Festinger, 1957) and self-perception (Bem, 1965) are reviewed in terms of the "insufficient justification" and "observer replication" laboratory paradigms. The controversy generated by their competing explanatory claims was evaluated in three separate contexts: as debate, theory, and metatheory. In a debating context it was concluded that Bem got the better of the controversy by observing the input requirements of his theory and marshalling against his critics evidence generated by their own failure to do likewise. Analytical and epistemological errors committed by the dissonance theory advocates were major factors in this conclusion. With respect to the more substantive context of theory-testing, it was concluded that Bem failed to establish the plausibility of the cognitive process postulated by the self-perception theory. A unique counter-instance was cited to demonstrate that self-perception is not a wholly viable alternative analysis of cognitive dissonance phenomena. In addition, an examination of Bem's adherence to a functional analysis in conjunction with a simulation methodology raised doubts that such a strategy could deliver the desired information concerning plausibility of the self-perception process. When viewed at the level of metatheory, however, Bem was considered to have had a substantial influence upon the working commitments of a small community of his colleagues. This conclusion was derived from a metaphorical application of Kuhn's (1962)

thesis concerning scientific revolutions to events in the recent history of Social Psychology. In this view, the self-perception theory is an historical marker which brings clearly into focus the transition of attitudinal research from a motivational-consistency "paradigm" to an information processing/attributional "paradigm".

Three experiments are reported which make use of Bem's credibility cueing procedure to articulate the newer "paradigm". The first experiment provides support for a fundamental hypothesis derived from the self-perception theory. Subjects' recall of a task was systematically influenced by external discriminative stimuli for self-credibility when internal memory cues were relatively weak, but not when such cues were relatively strong. The use of a statistic which takes into account subjects' differential guessing strategies increased confidence in the self-perception interpretation of these results. The second experiment attempted to extend the credibility cueing effect beyond the traditional impersonal cueing situation to one involving interpersonal discriminative stimuli for self-credibility. Although procedural insights rendered the results inconclusive, a serendipitous observation was made. The results suggested a novel hypothesis that different stimulus persons could have differential effects on subjects' self-credibility. A third experiment provided support for this hypothesis. When one live interviewer was manipulated as a discriminative stimulus for self-credibility, subjects' recall of a task was systematically influenced in accord with self-perception predictions. These effects did not occur

in the presence of a second live interviewer. Speculation was advanced concerning the psychological basis for differential credibility cueing properties of parties to social interactions with particular reference to the credibility cueing potential of police ~~interrogations~~.

ACKNOWLEDGEMENT

I would like to express my gratitude to Dr. Don Dutton who shared my initial ideas and tolerated my erratic pace with uncommon patience. In addition, I would like to thank Drs. Bob Knox, Tom Storm, John Yuille, and Mike Jackson for their helpful criticisms and suggestions, particularly in the preparation of the manuscript. I am indebted as well to Dr. Brian Little who indulged some of my more tenuous arguments. While he is in no way responsible for my excesses in the dissertation, he has contributed substantially to whatever clarity exists in the discussions relating to epistemology and metatheory. The characteristic enthusiasm with which Dr. Mike Humphreys followed the research and offered interpretive comment was also much appreciated.

Finally, I would like to acknowledge Clyde Curry, a fellow student whose insatiable appetite for table hockey contests delayed completion of the dissertation by several months. Marg, Andy, Schwartz, Elke and other "hut people" played similar roles to this same effect, for which I am eternally grateful.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
ACKNOWLEDGEMENT	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
OVERVIEW	1
CHAPTER ONE: THE SELF-PERCEPTION ALTERNATIVE TO COGNITIVE DISSONANCE THEORY	2
A. A Review of the Dissonance---Self-Perception Controversy	
The self-selecting subject "artifact"	
The self-selecting subject "fact"	
B. Some Evaluative Comments on the Controversy	
The analytical error	
The epistemological error	
CHAPTER TWO: THEORETICAL AND METATHEORETICAL CONSIDERATIONS	24
A. The Theoretical Contribution of Self-Perception	
Self-perception as an alternative analysis	
Functional analysis and simulation	
B. The Metatheoretical Contribution of Self-Perception	
Normal science and the paradigm	
Anomaly, crisis, and revolutionary science	
The resolution of scientific revolutions	
CHAPTER THREE: EMPIRICAL STUDIES IN THE INFORMATION PROCESSING/ATTRIBUTIONAL PARADIGM	56
Experiment 1	

TABLE OF CONTENTS (continued)

	Page
Experiment 2	
Experiment 3	
Credibility cueing in a non-laboratory context	
CHAPTER FOUR: SUMMARY AND DISCUSSION	125
REFERENCES	132

LIST OF TABLES

	Page
1. Mean proportions of hits, false alarms, and corresponding d' values for recall of word and trigram operations in each statement and light cue condition	71
2. Mean ratings of confidence in recall accuracy for word and trigram operations in each statement and light cue condition	71
3. Analysis of variance of recall scores (d')	72
4. Analysis of variance of confidence scores	73
5. Analysis of variance of uncorrected recall scores for words	77
6. Recall accuracy comparisons for trigram operations following false statements	80
7. Recall accuracy comparisons for trigram operations following true statements	80
8. Mean proportions of hits, false alarms, and corresponding d' values for recall in each statement and interviewer cue condition	88
9. Analysis of variance of recall scores	89
10. Mean proportions of hits, false alarms, and corresponding d' values for recall of word and trigram operations in each statement and interviewer cue condition for interviewer A	100
11. Mean proportions of hits, false alarms, and corresponding d' values for recall of word and trigram operations in each statement and interviewer cue condition for interviewer B	102
12. Analysis of variance of recall scores (d')	104

LIST OF FIGURES

	Page
1. Self-perception analysis of real and attributed causes in observer replication of the Festinger and Carlsmith (1959) experiment	21
2. Bem's simulation procedure (left) and a non-simulation approach (right) for strengthening his arguments	36
3. Recall accuracy predictions made by the self-perception theory	60
4. Recall accuracy for word and trigram operations in each statement and light cue condition	75
5. Recall accuracy in each statement and interviewer cue condition	90
6. Recall accuracy for word and trigram operations in each statement and interviewer cue condition for interviewer A	101
7. Recall accuracy for word and trigram operations in each statement and interviewer cue condition for interviewer B	103
8. Recall accuracy for words and trigrams in each interviewer identity condition	107
9. Recall accuracy following true and false statements in each interviewer cue and interviewer identity condition	108

OVERVIEW

The competing explanatory claims of Festinger's (1957) theory of cognitive dissonance and Bem's (1965) theory of self-perception have given rise to a lively controversy in Social Psychology. This dissertation takes a close look at the controversy, beginning with a review of the opposing positions and ending with an empirical test of some self-perception ideas.

Chapter One reviews a set of eight published articles in which the controversy between dissonance and self-perception theorists is represented. In an attempt to reach a clear understanding of the substance of their disagreement, the various arguments are evaluated within the narrow adversary framework of a debate.

Chapter Two examines the controversy more closely by evaluating the self-perception approach in the substantive contexts of theory-testing and contribution to metatheory.

Chapter Three describes three experiments which attempt to generate some necessary hard-headed support for the theory of self-perception and the new "paradigmatic" framework it reflects.

Chapter Four comprises a summary and discussion of the major accomplishments of the dissertation.

CHAPTER ONE: THE SELF-PERCEPTION ALTERNATIVE TO COGNITIVE

DISSONANCE THEORY

The recent history of attitudinal research in Social Psychology has been dominated by a loose network of ideas known to the faithful as the theory of cognitive dissonance (Festinger, 1957). According to this theory, an individual who is induced to behave in a manner that is inconsistent with his beliefs and attitudes will experience some mental discomfort as a result of holding the two contradictory cognitions, "I did X" and "I believe not X". The unpleasant state produced by this inconsistency (cognitive dissonance) is considered to have drive properties which motivate the individual to seek resolution (dissonance reduction) in a manner analagous to the reduction of such classical physiological drive states as hunger and thirst. It is presumed that this resolution will take the form of a mental manipulation whereby the cognitions are altered to produce a more harmonious state. Because beliefs and attitudes are subject to change whereas past behaviors are not, the cognition "I believe not X" should be more labile than the cognition "I did X". Consequently, attempts by the individual to resolve inconsistency, or to reduce dissonance, are more likely to be directed toward altering the cognition which has its foundation in a belief or attitude than the one which is derived from behavior. For example, the individual may convince himself that he actually holds the belief or attitude implied by his behavior. A change in his position from "I believe not X" in the direction of "I believe X" would be regarded as evidence for dissonance reduction.

Considerable research attention has been devoted to the notion of dissonance reduction as a vehicle for belief and attitude change. An experiment conducted by Festinger and Carlsmith (1959) has become a classic in the field and can be used to illustrate the laboratory paradigm for much of this work. In that study, college students performed a series of boring tasks as part of a psychology experiment (e.g., rotating pegs in a peg board), and were then asked to assist the experimenter in establishing an "expectancy" in the next subject. Some students were offered \$20.00 to tell the subject (an accomplice of the experimenter) that the tasks were fun and enjoyable, and others were offered \$1.00 to tell this same lie. A control group of students performed the same tasks but were not requested to tell the lie. According to cognitive dissonance theory, the \$20.00 people should experience little or no mental discomfort as a consequence of the two contradictory cognitions, "I performed boring tasks" and "I tried to convince someone that the the tasks would be fun and enjoyable". For these people, inconsistency can be averted by introducing a third cognition, "I told the lie because I was well paid to do so". The \$1.00 people, on the other hand, should experience some discomfort or dissonance because this additional cognition involving external justification for their behavior is not available to them. The theory predicts that these people will resolve the inconsistency by changing their beliefs about the tasks to a position which is more consistent with their behavioral descriptions of the tasks as fun and enjoyable. Although this experiment does not directly test the motivational

interpretation of how the changes are mediated, the results provide indirect support for this interpretation. Post-experimental ratings of how much they enjoyed the tasks revealed a significantly higher mean for subjects in the \$1.00 condition than for subjects in either the \$20.00 or control conditions.

In the past fifteen years a glut of cognitive dissonance research has appeared in print. This includes a great number of procedural and situational variations within the original insufficient justification laboratory paradigm, the introduction of slightly different laboratory paradigms, attempts to demonstrate and subsequently remove sources of confounding, a few non-laboratory investigations, and an occasional test of the proposed drive properties of cognitive dissonance. (For a review of representative dissonance hypothesizing and research in its hey day, see McGuire, 1966; for a modern "pop psychology" account of some of the same material, see Aronson, 1973.). In addition, dissonance research has traditionally attracted a number of critics who view the data as artifactual (see Rosenberg, 1963; Chapanis and Chapanis, 1964; Jordan, 1964; and Elms and Janis, 1965).

Recently, Daryl Bem (1965, 1967a, 1972) has taken a somewhat different critical approach by accepting the data supporting dissonance theory and proposing an alternative theory to explain them. He dismisses the motivational assumption as unnecessary, suggesting instead that subjects in "dissonance" situations infer, or discover, their beliefs and attitudes by discriminating the apparent controlling circumstances of their most recent relevant behavior. According to this self-perception

interpretation, many of an individual's self-descriptive statements (e.g., "I found those tasks fun and enjoyable.") which appear to be under the discriminative control of internal or private stimulation are, in fact, under the control of the same external or public events that others use in inferring that individual's inner states. Specifically, Bem argues that an individual's belief and attitude statements and the beliefs and attitudes that an outside observer would attribute to him are often functionally equivalent in that both sets of statements are inferences from the same evidence: the individual's overt behavior and the circumstances in which that behavior occurs. A frequently cited example involves the question, "Do you like brown bread?", and one reply that a person might make: "I guess I do, I'm always eating it.". This reply is considered to be functionally equivalent to the one an acquaintance might give for him: "I guess he does, he's always eating it.". Theoretically, the two persons are able to arrive at the same conclusion regarding the person's liking for brown bread because they use the same information in making their inferences. This information consists of the person's public eating behavior and its accompanying situational cues. In Bem's theory the individual is regarded as an observer of his own behavior and its controlling variables; accordingly, his belief and attitude statements are viewed as inferences which he makes from his own observations (For a Skinnerian analysis of the ontogeny of self-descriptive verbal statements, see Bem, 1967a.).

The Festinger and Carlsmith experiment can be reinterpreted in these terms. The self-perception theory considers the data from the

viewpoint of an outside observer who (1) hears a subject making favorable comments about the tasks to a fellow student, and (2) knows that the subject was paid \$20.00 (or \$1.00) to do so. If the observer is then asked to estimate the actual attitude of the subject he has heard, he is faced with an attribution problem: should he attribute the comments to some inner state of the subject (did he really feel that way about the tasks?), or should he attribute them to some causal factor in the environment (did something outside the subject make him say what he said)? If the observer had heard a subject making favorable comments for a large compensation (\$20.00), he is likely to make an attribution to the environment. That is, he may infer that the subject made his comments in order to obtain the money. Consequently, this subject does not appear credible to the observer in the sense that his behavior cannot be used by the observer as a guide for inferring an inner state or attitude toward the tasks. The observer's best guess, then, is to suppose that the subject's attitude is similar to that which would be expressed by anyone who was selected at random and asked for his attitude (i.e., the attitude of a control subject). On the other hand, if the observer had heard a subject making favorable comments for little compensation (\$1.00), he is likely to rule out the environmental attribution and make an attribution to the subject. That is, he may conclude that the money was insufficient to elicit favorable comments and therefore infer something about the subject's attitude ----- i.e., if the subject made favorable comments and was only paid \$1.00, then he must have really found the tasks to be fun and enjoyable.

The crucial premise of the self-perception theory is that subjects involved in the Festinger and Carlsmith experiment (and other "dissonance" experiments) are themselves behaving like hypothetical observers. Since this logic threatens to displace one of Social Psychology's most ubiquitous theories, it is not surprising that attempts at empirical validation have encountered a lively resistance. In the following review I have attempted to faithfully reproduce the substantive elements and Byzantine character of the exchange between dissonance and self-perception theorists.

A. A Review of the Dissonance—Self-Perception Controversy

According to Bem's analysis, self-perception is a special case of interpersonal perception. The basic propositions of the theory have been formally stated as follows:

Individuals come to "know" their own attitudes, emotions, and other internal states partially by inferring them from observations of their own overt behavior and/or the circumstances in which this behavior occurs. Thus, to the extent that internal cues are weak, ambiguous, or uninterpretable, the individual is functionally in the same position as an outside observer, an observer who must necessarily rely upon those same external cues to infer the individual's inner states.
(Bem, 1972, p. 2).

By postulating the functional equivalence of self- and interpersonal perception under limiting conditions of internal cue strength, Bem means that when these limiting conditions are satisfied, both self- and other descriptive verbal statements (the behavior that the theorist wishes to explain) are under the control of the same set of external, publicly observable variables (the behavior that the perceiver wishes

to explain). Assuming that "dissonance" situations do not violate its limiting conditions, one derivation from this analysis is that outsiders who observe a subject in a dissonance experiment should be able to accurately estimate that subject's attitude statement.

To test this derivation Bem (1967a) gave observer-subjects tape recorded descriptions of a college student who had participated in the Festinger and Carlsmith experiment. Control observers heard a detailed, non-evaluative description of the tasks and an outline of the alleged purpose of the experiment. In addition to this information, experimental observers were informed that the subject had accepted an offer of \$20.00 (or \$1.00) to tell the next subject that the tasks were fun. A brief conversation to this effect was also presented. When observers were asked to estimate an involved subject's attitude toward the tasks, Bem found that they were able to reproduce the inverse relationship between amount of compensation and subjects' attitude statements obtained in the original dissonance experiment. Using this same technique, which he originally termed "observer replication", Bem (1965, 1967a) claims two additional accomplishments. First, he has shown that observers can replicate a fairly wide range of "dissonance" phenomena, including opinion statements (Cohen, described in Brehm and Cohen, 1962, p. 73.), hunger judgments (Brehm and Crocker, described in Brehm and Cohen, 1962, pp. 133-136.), and toy ratings (Brehm and Cohen, 1959). Secondly, in an "extended replication" of the Festinger and Carlsmith experiment, he has shown that observers can reproduce some secondary patterns of data relating communication length, amount of

compensation, and attitude statement. The intricate form taken by these data need not be outlined here. Of importance for Bem is the observation that they are amenable to a self-perception interpretation, whereas dissonance theorists have either been silent or openly puzzled as to their explanation (see Brehm and Cohen, 1962, pp. 119-121). Bem concludes that these interpersonal replications support his analysis of self-perception as a viable alternative to cognitive dissonance formulations of attitudinal phenomena.

The self-selecting subject "artifact": Commenting upon Bem's analysis, Mills (1967) identifies what he considers to be a major flaw in all of the replication experiments. He refers to the Festinger and Carlsmith replication as an exemplar, and notes that Bem failed to inform his observers that the amount of money which an actual subject was paid to say the boring tasks were fun, was independent of his initial liking for the tasks. This failure could allow observers to successfully replicate the subject's attitude statement by adopting a judgment process which is presumed to differ from the one advanced by the self-perception theory. According to Mills' interpretation of Bem's position, Bem regards observers' differential estimates of subjects' liking for the tasks as evidence that observers have made the following assumption: the amount of money paid to a subject determined how much he liked the tasks. In effect, Mills presupposes that Bem thinks the causal relationship inferred by his observers is behavior + monetary justification → liking. The thrust of his critique is that observers could produce the same attitude statements by inferring the reverse

causal relationship, liking → monetary justification + behavior.

That is, observers could just as well have assumed that a subject's a priori liking for the tasks determined how much he was paid for his assistance (i.e., Because the tasks are boring, the experimenter is going to have to pay most people well in order to secure their participation. If, however, he finds some people who don't think the tasks will be boring, he won't have to pay them so much.). This is an implicit self-selecting subject hypothesis whereby observers assume differential initial attitudes on the part of subjects, which in turn determine the experimental condition to which they are allocated. In concluding this critique, Mills notes that Bem has not shown that observers can predict changes in subjects' attitudes in dissonance experiments.

In order to explore some implications of Mills' interpretation, Jones et al (1968) conducted an extensive series of experiments in which interpersonal observers attempted to replicate dissonance findings under varying conditions of initial information. In each experiment, observers were able to replicate successfully the attitudes or ratings of the original subjects when provided with Bem's (or Bem-like) descriptions. However, when other observers were given information about subjects' initial attitudes or ratings, a manipulation designed to rule out observers' use of a self-selecting subject hypothesis, they were unable to make similarly accurate estimates. Of course, had these latter observers been able to replicate successfully, the self-selecting subject explanation would have been rendered invalid. The failure to

replicate is taken as evidence that Bem's results were produced by the proposed judgmental artifact.

The self-selecting subject "fact": In a rejoinder to Mills, Bem (1967b) notes that this author is correct in asserting that observers fail to predict attitude change, but errs in assuming that this constitutes a valid criticism of his replications. Mills' error is attributed to a misunderstanding of the purpose of replications with respect to the phenomenology they are designed to duplicate. Bem contends that he has reproduced the original dissonance situations as seen by the subjects and observes that Mills has taken issue with him for not reproducing these situations as seen by the experimenters. Specifically, the replications were designed to duplicate the phenomenology of a subject who is asked to give an absolute judgment of his attitude following some behavior in which he has engaged. They were not designed to duplicate the phenomenology of an experimenter (or critic) who is interested in comparing a subject's present attitude with one he held prior to the experimental manipulation. A replication which duplicates this latter phenomenology, wherein attitude change is a psychological reality, would not be faithful to the theoretical goal of testing the functional similarity between interpersonal and self-perception.

The notion that initial attitudes are not salient information for subjects is implicit in Bem's dismissal of Mills' critique. A subsequent reply to Jones et al (Bem, 1968) clarifies and expands this position. Here again, Bem agrees with his critics and notes that their position is perfectly consistent with his own. He concedes that observers do

appear to attribute attitudes to subjects on the basis of a self-selecting subject hypothesis: "What must this person's attitude be if he is willing to behave in this fashion in this situation?". Jones et al are presumed to have erred, however, in assuming that this interpretation is different from the one advanced originally by Bem. Far from being an artifact of the replications, this judgment process is, according to the self-perception theory, the fact of the original experiments. In this view, the data of Jones et al is considered as support for Bem's original decision to regard subjects' initial attitudes as non-salient, and as evidence against the decision of Jones et al to regard these attitudes as salient. Speculating on the phenomenology of a subject in a dissonance experiment, Bem suggests that engaging in the behavior may provide such strong cues for the subject that any control exercised by the original attitude is swamped. In this connection he notes that Heider has made a similar observation in the context of interpersonal perception (see Heider, 1958, p. 54) and that the findings of at least two observer judgment studies are consistent with this idea (see Jones, 1966, and Jones and Harris, 1967). With respect to the replication methodology, the thrust of Bem's argument involves the possible levelling and sharpening effects of information given to observers by Jones et al. In their unsuccessful replications, the descriptions provided by these authors gave equal prominence to a subject's behavior and his initial attitude. Bem feels that this is a serious distortion of the situation facing an actual subject. In effect, behavioral information which is salient in a subject's phenomenology

has been made less salient for observers by Jones et al, whereas an initial attitude which is not salient information has been rendered salient.

In addition to their misunderstanding of his theory, Bem criticizes the Jones crew for failing to grasp the epistemological status of his interpersonal replications. He suggests that this confusion could have been avoided had he originally labelled his experiments "interpersonal simulations", a term which conveys their status as analogs to computer simulations. Stated in this way, the self-perception theory claims that an individual's attitude statements and an observer's judgments about them are "output statements" from the same "program". This hypothesis is tested by plugging in the interpersonal judgment program prescribed by the theory in place of the self-judgment process of a subject in a dissonance experiment. If the simulation succeeds, then it can be concluded that the process model embodied in the program is functionally equivalent to the process being simulated. It further implies that the selection of input statements (i.e., the omission of initial attitude information) was not in error. When their experiments are considered within this framework, the epistemological error committed by Jones et al becomes clear. In attempting to disconfirm Bem's program, they have adhered to the very methodology whose truth value is the subject of the attempted disconfirmation.

Piliavin et al (1969) suggest that the foregoing rejoinder is less than satisfactory. While conceding that the descriptions provided observers by Jones et al may well have rendered minimally relevant

information unduly salient, these authors point out that Bem has omitted potentially relevant information from his descriptions. This information originally appeared in the instructions of Cohen's essay-writing study and consisted of the experimenter (1) reminding a subject to take a strong stand against his own position, and (2) subsequently inviting the subject to consider the issue in the light of this stand. While these instructions presuppose, and draw attention to, a subject's initial attitude, they also seem to constitute an important bit of contextual information that would be a necessary input to Bem's judgment program. More importantly, they were an integral part of an original dissonance situation, and for this reason alone should be included as inputs to a simulation of that situation. According to this analysis, the outcome of a simulation which utilized a complete script of Cohen's original procedure would either provide strong support for, or do serious damage to, Bem's formulation. The findings parallel those of Jones et al. Observers who were provided with Bem's abbreviated descriptions were able to successfully replicate subjects' attitudes, while other observers given the complete descriptions were unable to do so. Piliavin et al conclude that Bem has produced the "right results for the wrong reasons". The wrong reasons consist of his choice of information for observers. Anticipating Bem's reply, that he has selected only salient information, they point out that it is then incumbent upon him to formulate, a priori, a general rule for selecting salient information.

In an attempt to put an end to this argument Bem and McConnell

(1970) conducted a non-simulation experiment to determine whether or not subjects in dissonance experiments "know" their initial attitudes after engaging in counter-attitudinal behavior. Subjects indicated their opinions on a campus issue and one week later wrote an essay arguing against those opinions. When asked for their final opinions, subjects in one experimental condition produced the usual dissonance effect ---- change away from initial opinion in the direction of the position advocated in the essay. The major results occurred in a second experimental condition. When asked to recall the opinions that they had expressed in the earlier session, these subjects reported opinions that closely paralleled the final opinions expressed by subjects in the other condition. In addition, this second group of subjects was also asked for their final opinions. It was found that their recall of initial opinions was more highly correlated with these final opinions than with the actual initial opinions themselves. The authors consider these findings as support for the self-perception analysis of a subject's post-manipulation phenomenology: data from the incoming behavior updates attitudinal information and destroys earlier information to the contrary.

The present investigation is also considered as a remedy for a weakness in the simulation methodology. Abelson (1968) has noted that in many social simulations it may be rather easy to obtain a good data fit by virtue of the small number of outcome variables of the model relative to the large number of parameters that can be juggled. This observation is particularly germane to Bem's cognitive simulations since

his outcomes merely involve the ordering of two means whereas the set from which his inputs can be drawn is sufficiently broad to generate controversy and demand for a selection rule. By returning to the original situation and demonstrating that the inputs of that situation and the inputs of the simulation are the same, Bem and McConnell have attempted to resolve this problem.

Finally, the authors concede that neither the simulations nor the present investigation provide a confrontation between self-perception theory and dissonance theory. For example, dissonance theory is not embarrassed by the finding that subjects fail to recall their initial opinions since forgetting of a conflicting opinion could itself be a mode of dissonance reduction. They conclude that a crucial, discriminating experiment is unlikely, and that preference for one theory or the other is "diminishing to a matter of loyalty or aesthetics". Although it may be a non sequitur, this comment is functionally significant in that it signals the end of the controversy. In the remainder of the chapter I shall discuss the prescriptive character of this signal.

B. Some Evaluative Comments on the Controversy

The acknowledgment of only eight papers in the preceding review invites the inference that I have arbitrarily limited the content of the controversy between dissonance and self-perception theorists. While this is not the case, the rationale for excluding additional papers did not extend beyond matters of semantics, style, and personal belief. Although several critics continue to attack his position, Bem has

ceased responding to them (at least in print). As a result we no longer have an exchange of arguments as a basis for controversy, nor do we have the careful (devious?) reformulations which added intellectual vigor to the exchange. In addition, the present author is firmly convinced that the controversy deserves to be put to rest on grounds that its extension would be epistemologically unsound. The remainder of this chapter constitutes a partial justification for that belief. It takes the form of a critical examination of the dissonance—self-perception controversy in the limited context of a debate. The main reason for viewing the controversy in this way is that it allows for an examination of the various arguments within the narrow adversary framework in which they have been exchanged. In this context we can consider what it is that the two sides have been arguing about and the rigor of their arguments. Like the protagonists themselves, we need not consider broader questions of what we feel they should be arguing about or the "truth" of propositions relating to such questions. Because the debate has proceeded primarily on Bem's own terms, we simply want to know if he has successfully defended the proposition that observers can replicate dissonance results when given appropriate descriptive information.

A consideration of the controversy in the context of a debate reveals two serious errors on the part of Bem's critics. By describing these errors I hope to show that Bem got the better of the debate. The errors can be traced to Mills' argument that observers could have employed a judgment process different from the one postulated by the self-perception

theory. In outlining an alternative process, Mills has introduced an analytical error which has been repeated by the other critics. In addition, by arguing that this alternative process was made plausible by Bem's failure to inform his observers about a subject's premanipulation attitude, Mills has invited the following inferences: (1) that provision of such information would prevent observers from replicating successfully, and (2) that an unsuccessful replication under these conditions would be evidence against the judgment process postulated by the self-perception theory and for the alternative process. By making this argument explicit and operational, Jones et al have committed an epistemological error.

The analytical error: Recall Mills' argument that observers could replicate dissonance results by assuming that an a priori attitude determined the amount of money a subject was paid for his behavior. This interpretation was offered as a plausible alternative to the self-perception analysis that behavior and monetary payment determined the attitude attributed to subjects by observers. Clearly these two interpretations pertain to different aspects of the judgment process. The self-perception theory is concerned with the real process by which a judgment is made, whereas Mills has addressed the issue of causal attributions which judges make subsequent to engaging in that process. To compare one with the other, as Mills has done, confuses an ethnoscience of the judgment process with a science of the process. While an ethnoscientific analysis would have real explanatory value in domains where the layman operates as a scientist, it more commonly furnishes insight into lay explanatory systems only. (For an entertaining

discussion of how scientists occasionally perceive the layman to operate like themselves, see Little, 1973, especially pp. 26-29.) Because self-perception is a domain in which laymen do not appear to behave very much like scientists, it becomes important to distinguish between lay causal attributions and the explanation offered by a theory. In this connection, Bem has acknowledged that "..... observers appear to attribute attitudes on the basis of an implicit self-selection hypothesis." (Bem, 1968, p. 270, Douglas' emphasis). In effect, observers make the popular assumption that behavior follows from attitudes by inferring that a subject held an attitude toward Festinger and Carlsmith's tasks before the experimenter solicited his assistance. They then deduce what that attitude must have been from observations of his behavior (he said the tasks were fun and enjoyable) and its context (payment of \$1.00/\$20.00). The self-perception theory assumes that attitudes follow from behavior and postulates that either (a) no such premanipulation attitude really existed (strong form of the argument), or (b) if a premanipulation attitude did exist, it is non-salient in the subject's post-manipulation phenomenology (weak form of the argument). In either case, the theory holds that a subject "discovers" his attitude only after he has produced some behavioral evidence from which, in conjunction with contextual cues, it can be deduced. The subject then assumes, incorrectly, that he held this attitude prior to assisting the experimenter, and that, in fact, both his decision to assist and the size of the payment offered were influenced by this attitude. Although I have discussed observers and subjects separately

here, the real judgmental dynamics postulated by the theory are the same for observers and subjects, as are their post-judgment attributions. This important distinction between real and attributed causes is shown schematically in Figure 1. Perhaps Mills' confusion in this regard stems from the present unusual context in which a lay explanatory system, the phenomenon of self-perception, is itself the subject of a scientific explanatory system, the theory of self-perception. Indeed, it is not all that surprising that such an error would be made in a discipline wherein "scientific" explanation does not often extend beyond common sense reasoning.

The epistemological error: The epistemological error committed by Jones et al is a curious one. It can be identified by comparing these authors' approach with the one taken by Bem. The self-perception theory holds that interpersonal judgment and self-judgment are functionally equivalent. A derivation from the self-perception analysis of cognitive dissonance phenomena is that observers should be able to reproduce attitude statements made by subjects in dissonance experiments.

According to the theory, a subject's initial attitude is not salient in his post-manipulation phenomenology and therefore should not be included in descriptive information given to observers who attempt to reproduce a subject's post-manipulation attitude statement. A successful replication under these conditions is viewed as support for the theory. Jones et al believe that interpersonal and self-judgment processes are not functionally equivalent in dissonance experiments. However, they have presented no evidence from dissonance research to support this

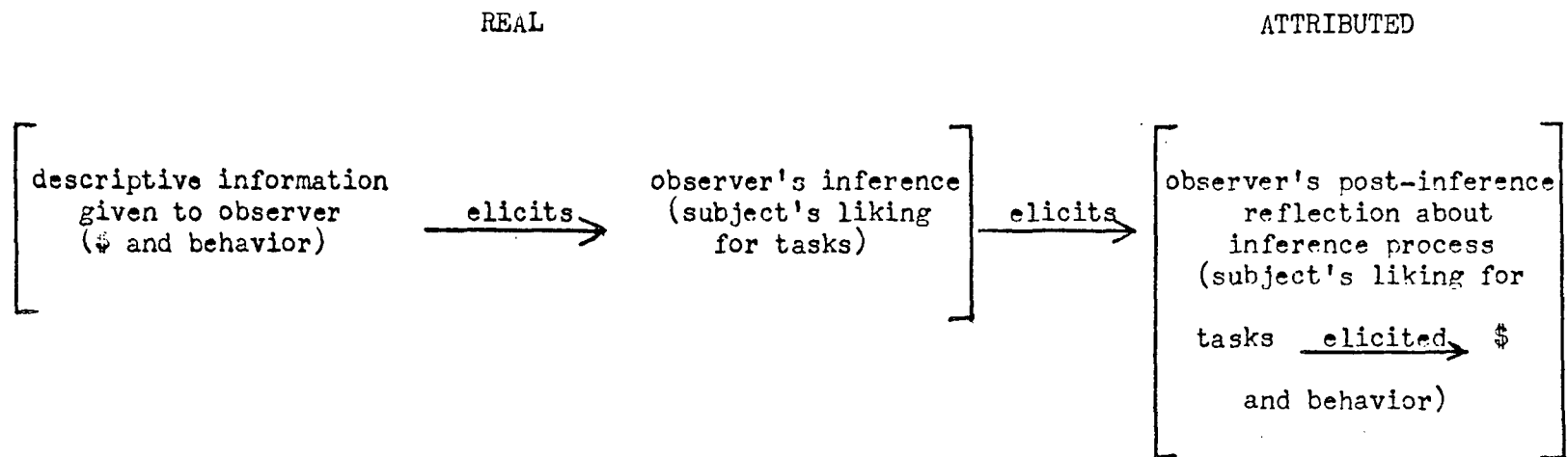


Figure 1. Self-perception analysis of real and attributed causes in observer replication of the Festinger and Carlsmith (1959) experiment.

belief. Instead, they have offered an armchair hypothesis that a subject's post-manipulation phenomenology contains salient information about his initial attitude, given this kind of information to observers and demonstrated that the replication fails under these conditions. In addition, they have reconfirmed that the replication succeeds when Bem's choice of descriptive information is provided. On the basis of these findings, Jones et al have concluded that the process embodied in Bem's replications is not a valid representation of the judgmental dynamics of a dissonance experiment.

This conclusion and its methodological context make little sense. If a theory indicates that certain information should be provided in order for observers to produce a particular result, and other information withheld because it is irrelevant, then it is unreasonable to conclude that the theory has been refuted when the provision of the irrelevant information washes out the results.¹ In a sense, a failure to obtain results under these conditions is consistent with the theory and should have been anticipated by Jones et al on a priori grounds. In this connection, Bem (1968) has pointed out that an attempted disconfirmation of his theory is bound to fail if it adheres to the very methodology whose truth value is the subject of the attempted disconfirmation. As to what would constitute an appropriate strategy for refuting the theory,

¹Freedman has made essentially the same point in a critique of role playing as a substitute for experimental methods (Freedman, 1969, pp. 112-113).

Abelson has noted:²

Ironically, what (Bem's) detractors should now really be doing if they must still simulate is to replicate (his) outcome with clearly bad descriptions to the observer, rather than to reverse (his) outcome with purportedly good descriptions.

Finally, in support of his claim that initial attitudes are non-salient, Bem has shown that subjects are unable to correctly recall these attitudes following dissonance manipulations. While this finding is subject to an interpretation other than the one offered by self-perception theory, the experimental procedure is significant in the present context. It represents the only non-simulation attempt to answer the question of input salience.

In short, Bem appears to have gotten the better of the debate by appealing to the input requirements of his own theory, the failure of simulations conducted with altered inputs, and non-simulation evidence of input salience. His critics' failure to state either theoretical or empirical grounds for altered inputs, their misunderstanding of Bem's theory and misuse of the simulation methodology, contributed in no small way to Bem's success.

²This comment was offered in a personal communication from Abelson to Bem and is reproduced in Bem and McConnell (1970). It was presumably made in reference to the question of "easy" data fits in simulations (re p. 15).

CHAPTER TWO: THEORETICAL AND METATHEORETICAL CONSIDERATIONS

This chapter extends the discussion of the dissonance---self-perception controversy to include substantive issues not considered in the previous context. First, Bem's approach is examined at the level of theory-testing in order to determine what it has revealed, and is likely to reveal, about a hypothetical cognitive process. Second, his approach is interpreted at the level of metatheory in an attempt to place it in the context of the contemporary history of attitudinal research in Social Psychology.

A. The Theoretical Contribution of Self-Perception

The self-perception theory was enunciated within a "radical behaviorist" framework. (For a general outline of the radical behaviorist approach, see Scriven, 1956; for an outline of this approach as it is applied to self-perception phenomena, see Bem, 1964, 1967a, and 1972.) While Bem's application of this strategy is articulate, it is open to question on one important point. A "true" radical behaviorist is one who insists that his analysis avoid reference to internal physiological or conceptual mediators, either real or hypothetical. Commenting upon his own theory, Bem notes:

It is probably the relaxation of this restriction which has robbed latter-day self-perception theory of its radical behaviorist flavor. One does not remain a behaviorist in good standing with repeated references to "inferential processes" and hypothetical inner dialogues ("What must my attitude be if I am willing to behave"). In order to reclaim membership, therefore, it should probably be said that such concessions to expository clarity do not, in my view, add anything to the explanatory power of the theory; it remains formally equivalent

to its earlier, albeit nearly incomprehensible, incarnation in the more rigid and arid vocabulary of radical behaviorism (Bem, 1964, p. 558). But a choice of language is not without heuristic consequences. For private "thinking" purposes, functional analysis remains my preference; but for expository purposes, English prose does not seem overly risky.
(Bem, 1972, p. 55).

This deft statement might best be viewed in the light of a bit of folk wisdom, which is itself "behaviorist" in nature: "if you want to know what someone is up to, watch what he does rather than what he says." In spite of his posture to the contrary, by speculating about subject phenomenologies and adopting a methodology analagous to computer simulations of cognitive processes, Bem appears to have a claim on a mediating process. By renouncing it, the claim is rendered obscure but no less real. In effect, the self-perception theory does not simply dispense with the motivational mediator postulated by the dissonance theorists, but instead replaces it with an information processing mediator. This issue is not raised as an objection, or even as a contradiction in terms, since the recognition of a mediating process need not detract from (nor as Bem has noted, add to) the accomplishments of a functional analysis. Without this clarification, however, the following discussion relating to a "cognitive process" would make little sense.

While Bem has successfully defended the proposition that observers can replicate dissonance results when given the descriptive information prescribed by his theory, it is not clear that he has shown anything about the importance of that information. Clearly, the information is important only insofar as it can be established that involved subjects could process it in the manner described by the theory. Hence, the crucial

issue here is not the suitability of informational inputs to a hypothetical cognitive process, but rather the plausibility of the process itself. Keeping strictly within Bem's own framework, this issue can be examined by considering two rather general questions. First, is there any evidence to suggest that self-perception is not a wholly viable alternative analysis of a cognitive process previously understood in terms of dissonance reduction? Second, is there any reason to believe that the strategy employed by Bem will deliver the information about the plausibility of the self-perception process that he desires? By looking at a "cognitive" cognitive dissonance phenomenon I hope to show that the answer to the first question is positive. By considering some of the implications of combining a functional analysis with a simulation methodology I hope to show that the second question has not been given sufficient attention. The intended purpose of this examination is to demonstrate that Bem has not established, in any "hard headed" way, the plausibility of the cognitive process postulated in his theory.

Self-perception as an alternative analysis: With respect to the question of whether or not self-perception is^a wholly viable alternative to dissonance analyses, it is important to keep in mind that the former theory was advanced as an alternative to the latter, and not the other way around. One consequence of this arrangement is that the self-perception theory should be embarrassed by data which can only be explained by dissonance theory, whereas the converse possibility ---- data explainable by self-perception theory only ---- has no implications for the question. The experimental manipulation of anticipated behavior is one non-trivial

phenomenon which merits attention in this connection.

The self-perception theory, and the simulation technique derived therefrom, require behavior as a prerequisite to any inferences that a person might draw concerning his or another's attitudes. Yet the notion of cognitive dissonance does not require that the behavior actually be carried out. The phenomenon of attitude change in situations where behavior is merely anticipated should, therefore, pose a problem for self-perception theory. To illustrate, Bem should have no difficulty interpreting Yaryan and Festinger's (1961) dissonance findings within the self-perception framework. In that experiment subjects who expended a great deal of effort in preparation for a possible future event (writing an aptitude test to be assigned to individuals at random) believed more strongly in the likelihood of that event than did subjects who engaged in a less effortful preparation. Seeing involved subjects diligently preparing for an event and others engaged in casual preparation should lead observers to make differential predictions about the subjective probabilities of the event for those subjects. Bem should, however, have difficulty accounting for the results of an extension of the Yaryan and Festinger experiment conducted by Arrowood and Ross (1966). These investigators found that subjects who merely anticipated expending effort in preparation for the possible event believed more strongly in the likelihood of the event than did subjects who anticipated a less effortful preparation.

In order to explore this possibility, Arrowood, Wood, and Ross (1970) repeated the Arrowood and Ross study using each subject as a

subject and observer simultaneously. They found that observers were unable to reproduce the subjective probability estimates of other involved subjects. Although this particular finding is equivocal due to artifacts in the Arrowood et al procedure, removal of the artifacts in a subsequent study by Dutton and Douglas (1972) revealed that observers were still unable to replicate successfully. While the epistemological status of this finding is unclear, it is not likely to be of comfort to Bem. On the one hand it could be argued that the anticipatory-effort-justification paradigm does not meet the initial requirements of Bem's theory on grounds that no behavior is involved. Because the paradigm lacks an informational input which is crucial to the self-perception process, no conceivable outcome of an experiment conducted therein could validate or invalidate the theory. The simple demonstration of altered subjective probability estimates after anticipated effort (Arrowood and Ross, 1966) is, then, sufficient to establish a phenomenon explainable by dissonance theory but not by the self-perception theory (in this view the Dutton and Douglas experiment is excess baggage). Curiously, the following expansion of a previously cited comment suggests that Bem might not be troubled by this conclusion:

At this juncture each theory appears capable of claiming some territory not claimed by the other, and one's choice of theory in areas of overlap is diminishing to a matter of loyalty or aesthetics.
(Bem and McConnell, 1970, p. 30).

This position fails to anticipate, or at least acknowledge, the possible significance of anticipatory-effort-justification phenomena for one's choice of theory. It is not only cavalier, but inappropriate when

considered within the present framework.

On the other hand, it could be argued that the anticipatory-effort-justification paradigm does, in fact, meet the initial requirements of Bem's theory. An indication that the theory could apply to situations where there is no obvious overt behavior comes from Bem's analysis of forced-compliance experiments in which subjects merely volunteer to engage in behavior (Bem, 1965, pp. 204-205). He contends that the commitment has crucial discriminative stimulus properties depending upon the incentive offered, and reports a successful simulation as evidence for this position.³ This analysis raises the possibility that by their presence in an anticipatory-effort situation subjects implicitly commit themselves to behavior, and that this commitment falls within Bem's framework. In this view the Dutton and Douglas study must be considered relevant. In particular, it is a rather unique and convincing unsuccessful simulation. It is unique because it avoids the problem of initial attitude salience by using extemporaneously generated objective probability estimates. It is convincing because the use of "initial attitude" in this fashion comes close to satisfying the demands of both Bem and his critics for a fair simulation of "dissonance" phenomena.

To summarize, if the anticipatory-effort-justification paradigm does not meet the initial requirements of Bem's theory, then Arrowood and Ross

³The role of commitment has not been systematically studied in these experiments. Bem, for example, would probably require an explicit statement of commitment on the part of a subject before observers could make accurate attitude estimates, whereas commitment in the Dutton and Douglas study was tacit in nature.

have demonstrated a "dissonance" phenomenon not explainable by self-perception theory. This demonstration has direct theoretical significance in that it clearly weakens the self-perception claim to alternative theoretical status. If, on the other hand, the anticipatory-effort-justification paradigm does meet these initial requirements, then the Dutton and Douglas simulation merits attention. Given that the self-perception alternative has been advanced primarily on the basis of simulation evidence, this particular simulation should indeed prove troublesome for Bem due to its unique character and unsuccessful outcome. Its epistemological status is indirect, or remote, however, since it is not clear just what simulations of the self-perception process, successful or otherwise, tell us about the process itself. This issue will be considered in response to the second of the two questions I have posed.

Functional analysis and simulation: The self-perception theory is a statement that an isomorphism exists between the subject area of interpersonal perception and that of self-perception. By specifying limiting conditions of internal cue strength the theory indicates that this isomorphism is incomplete. It is considered to hold for "dissonance" situations, however, and empirical tests of this application have relied primarily upon a simulation methodology. While conceding that simulations of self-perception phenomena do not constitute an adequate substitute for an intensive study of the phenomena themselves, Bem feels that this methodology is useful as a "plausibility test" (Bem, 1968, p. 273; Bem and McConnell, 1970, p. 25). Let us consider some of the implications of this kind of testing.

According to the self-perception theory, when it comes to "knowing" his own attitudes "(an) individual is functionally in the same position as an outside observer." (Bem, 1972, p. 2, Douglas' emphases). In a comment upon the epistemological status of his simulations, Bem (1968, p. 273, Douglas' emphases) makes the following statement: "(a) successful simulation implies the same thing that a successful computer simulation implies, namely, that the process model embodied in the "program" is functionally equivalent to the process being simulated"

A distinction between the theoretical and epistemological uses of the term "functional equivalence" is rather important. I take "functional equivalence of position" to mean that both self- and other-descriptive attitude statements are under the control of the same set of publicly observable variables. In keeping with my earlier argument concerning Bem's claim to a mediating process (p. 25), this interpretation in turn implies that self- and interpersonal inference processes are structurally equivalent. Indeed, a more conservative interpretation would render the theory comparatively trivial and uninteresting. I take "functional equivalence of process" to mean that an interpersonal perception process is sufficient to generate the outputs of a self-perception situation. It does not imply that the two processes are structurally equivalent, but instead leaves open the possibility that different processes mediate the same outcomes. This is an obvious limitation of the simulation methodology. Since the main theoretical issue concerns functional equivalence of process only insofar as evidence to this effect influences the plausibility of an inference regarding structural equivalence of

process, additional problems are posed by this limitation. For example, it would be important to determine (1) how the simulation methodology can be used and/or strengthened to support a theory which implicitly postulates the structural equivalence of cognitive processes, and (2) the extent to which this has been accomplished by Bem.

A necessary first step in applying the simulation methodology is to demonstrate functional equivalence of process between the referent and "model" situations (the areas of self-perception and interpersonal perception in this case). A demonstration of both outcome equivalence and input equivalence accomplishes this objective. While there can be little doubt that Bem has demonstrated outcome equivalence, his claim concerning input equivalence is less secure. To reiterate, Bem has argued that the similarity in outcomes between dissonance experiments and his simulations indicates that (1) the process involved in the simulation is functionally equivalent to the process of the dissonance experiment, and (2) that the inputs to the simulation must be the same as the inputs to the dissonance experiment. While his critics dispute this claim on both counts, it was concluded that Bem got the better of the argument in a debating context. As is often the case in debates, however, the winning arguments fail to clarify important underlying issues. Note that the simulation equation consists of three elements ----- input sameness, output sameness, and process sameness ----- and that an inference regarding any one requires evidence about the remaining two. Although Bem has provided empirical evidence on the output side, his evidence on the input side is inferential and tenuous. His argument runs as follows: particular

inputs are suggested by the self-perception theory, and the successful matching of dissonance outcomes when these inputs are employed is evidence for the proper selection of inputs; a failure to match outcomes using different inputs is evidence that these inputs were inappropriate. This is a clear attempt to pull a theory up by its own bootstraps since the analysis requires the application of the very theory that is to be supported by the evidence in question.⁴ Penner and Patten (1970) make the same point in noting that Bem's use of the term "functionally equivalent" means, in effect, that different mediating processes and/or different inputs produce the same outcomes. They suggest that in trying to decide whether or not inputs or processes are the same, some evidence may be derived from the fact that they lead to the same outcome. However, it would be desirable to have assurance of process as well as outcome sameness in an attempt to establish input sameness, and of input as well as outcome sameness in an attempt to establish process sameness.

The thrust of the preceding argument is that, in the absence of an acceptable demonstration of input equivalence, Bem's simulations do not constitute evidence for the functional equivalence of self- and interpersonal perception processes. Consequently, an inference regarding structural equivalence would not be plausible. However, since an "acceptable" demonstration of input equivalence would require direct evidence of the very process which is the subject of investigation, this represents an

⁴Dissonance theorists occasionally employ a similar tactic when confronted with negative results. When their dependent measures do not show the predicted attitude changes, this is taken as evidence that their experimental manipulations failed to arouse dissonance.

impossibly stringent criterion for those who would continue to simulate. A necessary second step in applying the simulation methodology, then, is to accept the notion of functional process equivalence more or less on faith and consider ways of strengthening the simulation arguments. In this connection, Bem feels that his simulation arguments have been strengthened by "an extended replication of some secondary data patterns" in the Festinger and Carlsmith experiment (Bem, 1967a), and also by a demonstration of input equivalence between a dissonance situation and a simulation of that situation (Bem and McConnell, 1970). These demonstrations are not convincing. One means of strengthening a simulation argument is "to design the simulation so as to generate as large a number of outcome variables as possible. The more outcomes that can be validated, the merrier ——— and the more convincing the underlying theory." (Abelson, 1968, p. 344). While it is quite plausible that a single (type of) successful outcome match could be the product of a singular (type of) input artifact, it is considerably less plausible that a large and varied set of outcome matches could be produced by either a singular (type of) input artifact or by a large and varied set of input artifacts. Bem's replications of dissonance experiments, "extended" or otherwise, do not go very far towards ruling out artifacts by generating the desired large and varied set of matchable outcomes. In fact, by restricting his simulations to self-perception phenomena traditionally interpreted within the relatively narrow domain of dissonance theory, Bem has pretty much precluded this possibility. The second demonstration is unconvincing because it is open to the dissonance interpretation outlined earlier (p. 16).

By taking a non-simulation approach which simply mirrors Bem's procedure, a third way of strengthening his arguments is suggested. Although his terminology implies that he has adopted interpersonal perception as a model for his referent process, self-perception, in fact Bem has used modeling as a metaphor rather than as a faithful description of an explanatory tool (For an understanding of the explanatory use of modeling strategies in the social sciences see Anderson and Moore (1966) and Harré and Secord (1972); for a classic application of this approach to attitudinal research see McGuire's immunization model for propaganda defense (1964)). In outlining Bem's procedure and the non-simulation approach suggested here, this small concession to proper terminology is reflected in the use of the term "subject area" in place of "referent" and "equivalent area" in place of "model".

The simulation paradigm is outlined on the left-hand side of Figure 2. It consists of plugging subject area inputs (a') into the equivalent area process (b) and demonstrating that the outputs of that process (c) match those of the subject area process (c'). A non-simulation paradigm that could be used to strengthen (or weaken) arguments based upon this procedure is outlined on the right-hand side of Figure 2. It consists of giving inputs from the equivalent area (a) to the subject area process (b') in the hope of demonstrating that the outputs of that process (c') match those of the equivalent area process (c). A potentially useful framework within which this demonstration could be attempted is offered by the "correspondent inference" theory of Jones and Davis (1965). This is a theory of interpersonal perception which is clearly formalized and

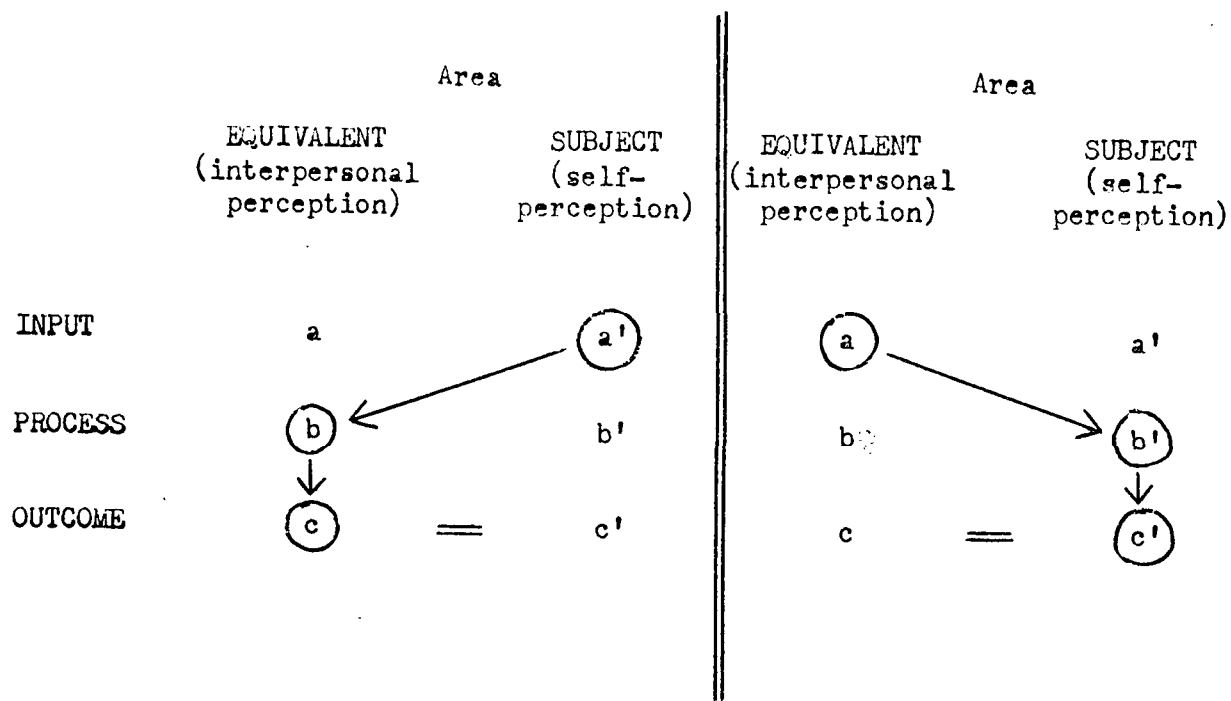


Figure 2. Bem's simulation procedure (left) and a non-simulation approach for strengthening his arguments (right).

empirically supported. Moreover, the theory deals with inferences from the same kind of discrete public events that are of interest to Bem.⁵ If individuals' self-inferences were found to obey the same laws of correspondence that apply to observers' inferences (i.e., same inputs produce same outputs), this would be evidence for functional process equivalence. Because the correspondent inference theory applies to a much broader set of situations than does the theory of cognitive dissonance, it could conceivably generate a sufficiently large number of outcome variables as to make plausible inferences about structural process equivalence.

I am unaware of any systematic application of this particular research strategy. Certainly Bem has not attempted to strengthen his simulation arguments in this way. There is, however, a growing body of empirical literature which compares individual actors' perceptions of the causes of their own behavior with outside observers' perceptions of those causes. Recently Jones and Nisbett (1972) reviewed much of this work and concluded that individuals have a pervasive tendency to attribute their own actions to situational requirements, whereas observers tend to attribute the same actions to personal dispositions of the actor. Of particular relevance to Bem's theory is their premise that actors and observers differ fundamentally in their processing of the same information.

⁵By way of comparison, other interpersonal perception theories such as Kelley's (1967) deal with inferences from events extended in time. Although corresponding concepts could be abstracted from such theories, they could not be integrated into the present non-simulation paradigm as easily as could those from Jones and Davis' theory.

This raises the possibility that the outcome of a systematic application of the research strategy outlined above would weaken rather than strengthen Bem's simulation arguments.

To conclude, Bem's claim concerning functional process equivalence seems overly optimistic. As a result, his argument for abandoning the simulation methodology in favor of an intensive study of self-perception phenomena per se becomes especially appropriate. Abelson, an author whose views are occasionally cited by Bem, suggests a similar conclusion with his definition of a simulation:

Simulation is the exercise of a flexible imitation of processes and outcomes for the purpose of clarifying or explaining underlying mechanisms involved. The feat of imitation per se is not the important feature of simulations, but rather that successful simulation may publicly reveal the essence of the object being simulated.
(Abelson, 1968, p. 275).

This goal is not reflected in the simulations carried out by Bem (and his critics). In fact, the major empirical accomplishment of this particular application of the simulation strategy is the feat of outcome imitation. Although evidence to the effect that individuals could process information in the manner described by the self-perception theory might be forthcoming from a more rigorous application of this approach, contrary evidence seems a no less likely possibility. As a consequence, further entertainment of the crucial notion of structural process equivalence would seem to require considerable imagination on the part of those who wish to go on simulating.

B. The Metatheoretical Contribution of Self-Perception

If theoretical disagreements were always settled by the kind of logical argument and experimentation outlined in the preceding section, it could be properly concluded that Bem has made no substantive contribution to the study of attitudinal phenomena. In fact, given that type I errors are especially loathsome to scientists, it might even be argued that he has done a disservice by proclaiming a relationship among variables which he has failed to support. If, however, theoretical disagreements are sometimes decided by "mystical" factors which are psychological and historical rather than logical and empirical, then it might be properly concluded that Bem has made a fairly significant contribution, one which extends well beyond the scope of attitude research. In the remainder of this chapter Bem's approach is considered at the level of metatheory, in terms of its real and potential effects on the informal rules that guide the activity of a community of social psychologists. These considerations make use of a loose interpretation of Thomas Kuhn's (1962) thesis concerning the structure of revolutions in the natural sciences.

Hopefully, confusion can be averted by emphasizing at the outset that Kuhn's thesis is adopted here in a sense similar to Bem's use of the modeling metaphor described earlier. In this context the arguments are not considered to be "true" in a strict sense, but rather they are thought of as useful in explicating a subject which would otherwise be misrepresented. For example, when his critics persisted in misinterpreting his "observer replications" as purported demonstrations of self-perception principles, Bem introduced the notions of "modeling" and "simulation" in

an attempt to clarify their lesser status as plausibility tests. In a related way I am introducing Kuhn's thesis as a basis, or vantage point, for viewing an event whose fundamentally non-scientific character is traditionally misrepresented in favor of explanation in terms of scientific enterprise alone. Granted that social psychologists do not practice normal science in a Kuhnian sense, do not embrace universal paradigms, and hence do not undergo periods of revolutionary science, still they do experience activities like these and for this reason it is important to recognize the role of such experience in shaping a part of their discipline. Perhaps the major danger in applying Kuhn's arguments to Bem's work is one of pretension, since the concepts themselves suggest an influence on a scale which is much too grand for the events at hand. Yet, the psychological dynamics denoted by these same concepts applies quite well to these events. If it is kept in mind that the analysis is applied metaphorically to a relatively small community of academics with restricted conceptual interests, the connotation of grandeur might be avoided. As a colleague has observed in a very apt euphemism, "Bem's contribution is hardly Michelson and Morley stuff."; on the other hand, when the working commitments of one's own academic community shift in a way that is essentially captured by a Kuhnian kind of analysis, this shift is worth bringing into focus ----- even in a PhD thesis.

Normal science and the paradigm: "Normal science" is the term Kuhn uses to describe the activity in which most scientists are engaged almost all of the time. Its fundamental characteristic is that it takes place within a strong network of commitments shared by a large scientific

community. These commitments constitute a paradigm which describes what a portion of the world is like, specifies which problems are important and what constitute adequate solutions, and provides a general set of rules as to how one ought to proceed in order to make further scientific contributions. When a scientific community acquires a paradigm it acquires a criterion for choosing problems that its members will be encouraged to undertake. A proper interpretation of this term implies that Psychology is so fundamentally different from Physics, Chemistry, or Astronomy, that it more closely resembles a random fact-gathering exercise than a science. In particular, there are no parallels in Psychology to the universally recognized achievements of Newton, Lavosier, or Copernicus which for a time provided model problems and solutions for their respective communities of practitioners. In a strict comparative sense, Psychology is pre-paradigmatic and therefore a prescientific discipline.

Although this argument can be advanced even more forcefully with respect to an individual subject area like Social Psychology, its significance is largely metaphysical. Most social psychologists think of themselves and each other as scientists. Moreover, the bulk of their professional activity takes place in accord with the kinds of shared commitments that define a paradigm. Although I shall retain the term "paradigm" as a label for these commitments, since they are even more circumscribed than the kinds of fundamental traditions that psychologists might recognize as paradigms (i.e., behaviorist, humanist, psychoanalytic) it is more appropriate to think of them as "miniparadigms". These differ from paradigms in the natural sciences primarily in the extent

of their membership, normative influence, and precision. For example, at any one time all physicists share the same paradigm, whereas all social psychologists do not. Moreover, physicists rarely, if ever, think about possible alternatives to a current paradigm, whereas their actual existence is usually a matter of toleration or indifference for social psychologists. Finally, problems and solutions are clearly specified in Physics so that there is little interpretive latitude, whereas Social Psychology leaves considerable room for interpretation in both matters.

When paradigms are considered in terms of the kind of normal scientific activity which they sanction, this loose paradigmatic view of Social Psychology seems quite justified. From this perspective a definite parallel can be seen between natural science and the work of most social psychologists. Because it is contrary to the popular view of science, however, this parallel is non-obvious. It is based on the observation that a paradigm does not aim to produce major substantive novelties. A basic requirement of a problem in natural science is that it be solvable, that is, that it have a known solution. In fact, the outcome must be so clearly anticipated that a failure to approximate it can be, and usually is, seen as a failure on the scientist's part. This seemingly peculiar arrangement suggests an intriguing question: why are such problems undertaken? By relating normal science to a puzzle-solving metaphor, Kuhn suggests the answer. Puzzles are sets of problems that can serve to test ingenuity or skill in solution. While their outcome can be anticipated in detail, the way to achieve that outcome

is very much in doubt. In Kuhn's view, this is where the interest lies in doing research in the natural sciences:

(It involves) achieving the anticipated in a new way, and it requires the solution of all sorts of complex instrumental, conceptual, and mathematical puzzles. The man who succeeds proves himself an expert puzzle-solver, and the challenge of the puzzle is an important part of what usually drives him on. What challenges him is the conviction that, if only he is skilful enough, he will succeed in solving a puzzle that no one before has solved or solved so well. Many of the greatest scientific minds have devoted all of their professional attention to demanding puzzles of this sort. On most occasions any particular field of specialization offers nothing else to do, a fact that makes it no less fascinating for the proper sort of addict.
(Kuhn, 1962, pp. 36-38).

The social psychological parallel to this preoccupation is implicit in the strikingly similar metaphor which William McGuire has chosen to describe research activity in his discipline (McGuire is an unusual social psychologist who shares Kuhn's appreciation for the psychology and history of his own profession). Although his view is intentionally pejorative, whereas Kuhn's is flatly descriptive, it captures the fundamental quality and fascination of puzzle-solving for social psychologists:

Experiments naturally turn out to be more like demonstrations than tests. If the experiment does not come out "right", then the researcher does not say that the hypothesis is wrong but rather that something was wrong with the experiment, and he corrects and revises it, perhaps by using more appropriate subjects, by strengthening the independent variable manipulation, by blocking off extraneous response possibilities, or by setting up a more appropriate context, etc. Sometimes he may have such continuous bad luck that he finally gives up the demonstration because the phenomenon proves to be so elusive as to be beyond his ability to demonstrate. The more persistent of us typically manage at last to get control of the experimental situation so that we can reliably demonstrate the hypothesized relationship.

But note that what the experiment tests is not whether the hypothesis is true but rather whether the experimenter is a sufficiently ingenious stage manager as to produce in the laboratory conditions which demonstrate that an obviously true hypothesis is correct. In our graduate programs in social psychology, we try to train people who are good enough stage managers so that they can create in the laboratory simulations of realities in which the obvious correctness of our hypothesis can be demonstrated.
(McGuire, 1973, p. 449).

In addition to their prescriptive function in identifying appropriate puzzles, social psychological paradigms play an important supportive role by insuring that such metaphors will neither occur to the faithful, nor be taken too seriously by them.

From the mid-1950's until quite recently, appropriate puzzles for attitude change researchers have been identified mainly by a single paradigm. This paradigm has committed a substantial community of social psychologists to the proposition that psychologically inconsistent cognitions are a significant motivational basis for attitude change. This proposition has been formally expressed in a half dozen similar theories, variously termed "balance" (Heider, 1946, 1958), "symmetry" (Newcomb, 1953), "congruity" (Osgood and Tannenbaum, 1955), "psychologic" (Abelson and Rosenberg, 1958), "cognitive dissonance" (Festinger, 1957), and "affective-cognitive consistency" (Rosenberg, 1960). Because the cognitive dissonance theory has attracted by far the largest and most enduring group of practitioners, dissonance research can be properly viewed as the model expression of the activity sanctioned by the motivational-consistency paradigm. In fact, no other social psychological theory has spawned a tradition of professional activity more characteristic of the metaphorical descriptions of normal research given by McGuire and

Kuhn. (For what amount to detailed applications of the stage management metaphor to dissonance research, see McGuire, 1969, pp. 139-140; Kelley, 1967, pp. 224-231; and Elms, 1972, pp. 210-219.)

Anomaly, crisis, and revolutionary science: When normal science is successful, it is a highly cumulative exercise which does not produce empirical or theoretical surprises. Those unsuspected phenomena that do appear are handled by minor adjustments of theory. On occasion, however, normal research inadvertently leads to discoveries which demand adjustment on a much larger scale. Because their resolution is not conceivable within the existing paradigm, these discoveries are more than just puzzles. They are major anomalies which signal a breakdown in normal science. Recognition of this breakdown by particular members of the scientific community marks the beginning of a period of special activity in which a new and significantly different set of rules will be elaborated to assimilate the anomaly. Kuhn uses the term "revolutionary science" to describe this period of special activity.

Although precipitated by events that are fundamentally scientific, the definitive aspect of revolutionary science is that it is not really scientific at all. Because the object of normal science is the solution of puzzles for whose very existence the validity of the paradigm must be assumed, the effects of this activity are limited. Normal science can articulate a paradigm and ultimately lead to the recognition of major anomalies, but it cannot correct a paradigm. Instead, this failure corrects itself by precipitating a crisis state in which the paradigm's validity is questioned and the rules of normal science are

loosened. In effect, the failure of the paradigm legitimizes the kinds of fiddling and tinkering activities characteristic of the popular stereotype of the scientist. Because the paradigm no longer provides satisfactory puzzles, he tries experiments just to see what will happen and generates speculative theories. In addition, this approach is often accompanied by a search for rules and assumptions in a philosophical analysis of the contemporary research tradition. Sometimes these unusual procedures provide the incremental data needed to establish a new paradigm. However, Kuhn contends that more often the new paradigm will emerge all at once, from a relatively sudden and unstructured perceptual "experience" which enables a scientist to see a problem in a new way that for the first time permits its solution. This experience is likened to the pioneering gestalt demonstrations in which a subject fitted with inverting lenses undergoes a transformation of vision. Initially his perceptual apparatus functions as it had been trained to function and he sees the world upside down. After a period of extreme disorientation and personal crisis, he begins to learn to deal with his new world and his entire visual field flips over. The assimilation of a previously anomalous visual field has reacted upon and changed the field itself. This switch of gestalt is prototypical of a paradigm shift. (For an account of the original gestalt demonstrations see Stratton, 1897; for a contemporary metaphor which captures the proposed epistemological transformation as well as the perceptual alteration, see the sociological works of Castenada, 1963, 1971, and 1972.) The eventual transference of commitments to the new paradigm by the scientific community, and the controversy which accompanies

this transition, is a scientific revolution.

Because Social Psychology is practiced within a number of mini-paradigms rather than within a single pervasive paradigm, the discipline is not subject to the same kind of revolutionary paradigm shifts that can be identified in the natural sciences. That is, there are no faithful parallels to the shifts from Aristotelian to Newtonian dynamics, or from Newtonian to Einsteinian theory. However, shifts from one miniparadigm to another have occurred in the sense that different sets of commitments have dominated social psychological communities at particular times. For example, before Festinger and the motivational-consistency people took over, the study of attitude change was dominated by a loose reinforcement tradition headed by Carl Hovland and his associates at Yale. What is central to the Kuhnian thesis that I am developing here is an impression that the Festingerian tradition is now giving way to an information processing/attributional approach to the study of attitudinal phenomena. Each of these three approaches is not only sufficiently articulated in theory, method, and style to constitute a paradigmatic framework in the restricted sense in which I am using the term, but also sufficiently distinct to justify using the term "paradigm shift" to characterize the transference of academic allegiances from one to another.

Although the dynamics of a paradigm shift in natural science are non-scientific and often psychological, it was noted earlier that events which precipitate the shift are fundamentally scientific. They involve the production of anomalies by normal scientific activity. In partial

contrast, both the dynamics and precipitating factors of paradigm shifts in Social Psychology tend to be of a psychological nature. Like the natural scientist who spends almost all of his time demonstrating that the world fits into the conceptual boxes provided by his paradigm, a similar activity preoccupies most social psychologists. Whereas the natural scientist's paradigm tends to be sophisticated and esoteric, the social psychologist's paradigm contains a strong element of common sense in the form of folk psychological principles. Research in the Hovland tradition (e.g., trustworthy communicators elicit greater attitude changes than do untrustworthy ones) and interpersonal attraction studies (e.g., people like others who possess attitudes similar to their own) are examples.⁶ Because a failure to demonstrate a folk psychological principle is almost always attributed to the psychologist rather than to nature, this standard brand of normal Social Psychology virtually guarantees the absence of anomalies. It is not surprising then, that a flashier brand of Social Psychology flourishes against this austere background. Its fundamental feature is the demonstration of non-obvious or counterintuitive principles which, by definition, guarantees the emergence of anomalies. Research on bystander intervention (e.g., within limits, the more witnesses to an emergency the less likely a victim will receive aid) and interpersonal attraction studies (e.g., a highly attractive target is judged to be even more attractive when he commits a faux pas) are examples. (For a discussion of this brand of research

⁶ Occasionally this style of Social Psychology is raised to a non-trivial level by an insightful practitioner such as Fritz Heider.

and the values associated with it, see the debate between Kenneth Ring (1967) and William McGuire (1967).) Because its products will sometimes be anomalous in more than just a counterintuitive sense, the potential for revolutionary change in Social Psychology lies in this activity. By articulating in theory those special anomalies that directly contradict or threaten to replace an existing paradigm, a particularly creative practitioner can precipitate a crisis among a small community of his fellows. A subsequent rejection of the time-honored way of doing things in favor of the newer approach, together with the controversy that almost always accompanies such a transition, would define a scientific revolution for that community.

According to this analysis, the dynamics of paradigm shifts in Social Psychology are quite different from those in the natural sciences. In the natural sciences a new paradigm emerges in response to a crisis initiated by a failure of the existing paradigm. This paradigm shift is mediated by a loosening of the rules of normal science. In Social Psychology it is the prior existence of a new candidate with the aforementioned characteristics that provokes a crisis for an existing paradigm. A loosening of the abstract "rules" of normal Social Psychology is a prerequisite for this paradigm shift. Although the conceptual, methodological, and empirical precedents for doing the standard brand of normal Social Psychology are taken for granted by most practitioners, a general scepticism concerning at least some of these precedents is a trademark of the parallel, flashier brand of Social Psychology. It is this prevailing scepticism which favors the kinds of speculating

and fiddling activities that ultimately enable some of these latter practitioners to see their world in new and different ways.

The emergence of both cognitive dissonance and self-perception theories can be viewed in terms of this analysis. Attitude change research conducted in the reinforcement paradigm did not produce the kinds of anomalies which would result in the breakdown of that tradition. Instead, a breakdown was precipitated by speculation and data which were counterintuitive and counterparadigmatic. When the reinforcement paradigm was confronted with the incentive effects predicted by dissonance theory, a period of protracted controversy marked the shifting of a substantial body of professional allegiance away from that paradigm. Subsequent efforts were directed toward articulating (and exploiting) the newer motivational-consistency paradigm in general, and the theory of cognitive dissonance in particular. In a parallel fashion, the dissonance tradition has never been seriously threatened by the work of dissonance researchers. It has been threatened by a counterintuitive hypothesis and evidence to the effect that individuals do not have privileged access to their own internal states. In terms of Kuhn's analysis, it is significant that Bem acknowledges Ryle's (1949) philosophical analysis of the concept of mind as a primary source for this hypothesis.⁷ By advancing it as a formal alternative to the dissonance formulations, he has made explicit the counterparadigmatic property of the self-perception hypothesis. The

⁷ Additional sources cited by Bem are also far removed from the contemporary research tradition of Social Psychology. These include Mead (1934), Skinner (1957), and Chappell (1962).

ensuing controversy in which the two sides talked through each other, each refusing to grant the other's non-empirical assumptions, is typical of the kind of exchange that occurs between scientists who view their world from the perspective of different paradigms. Moreover, it is indicative of the kind of activity which precedes the transition from one paradigm to another.

The resolution of scientific revolutions: What causes a scientific community to abandon an old paradigm in favor of a new one? Ostensibly, it is the recognition that the new paradigm fits the facts better than the old one. "Critical tests" between competing paradigms play a major role in this view. Results favoring a new paradigm are presumed to be especially persuasive. Kuhn suggests that this formulation is a distortion which makes the task of choosing between paradigms appear easier and more straightforward than it actually is. The simple fact that a new paradigmatic candidate is available implies that there is more than one way of viewing the same scientific problems, and more than one set of standards for defining solutions. This insures that the proponents of competing paradigms will be at least slightly at cross purposes. The earlier observation that Bem and his critics have been unwilling to grant the non-empirical assumptions that the other needs to make his case is an example of this incommensurability of standards. Hence we come back to Kuhn's notion of a gestalt shift:

In a sense that I am unable to explicate further, the proponents of competing paradigms practice their trades in different worlds. One contains constrained bodies that fall slowly, the other pendulums that repeat their motions again and again. In one, solutions are compounds, in the other mixtures. One is embedded

in a flat, the other in a curved, matrix of space. Practicing in different worlds, the two groups of scientists see different things when they look from the same point in the same direction. Again, that is not to say that they can see anything they please. Both are looking at the world, and what they look at has not changed. But in some areas they see different things, and they see them in different relations one to the other. That is why a law that cannot even be demonstrated to one group of scientists may occasionally seem intuitively obvious to another. Equally, it is why, before they can hope to communicate fully, one group or the other must experience the conversion that we have been calling a paradigm shift. Just because it is a transition between incommensurables, the transition between competing paradigms cannot be made a step at a time, forced by logic and neutral experience. Like the gestalt switch, it must occur all at once (though not necessarily in an instant) or not at all.
(Kuhn, 1962, p. 150).

How, then, is this shift induced in a scientist? Kuhn suggests that often it is not induced at all. Instead, a new paradigm is adopted simply because its opponents die out while a new generation is growing up to be familiar with it. To the extent that a paradigm shift can be effectively induced by argument, Kuhn identifies two particularly persuasive claims that might be made for a new paradigm. First, if it can be shown that the new paradigm solves the problem(s) that led the old one to a crisis, then many scientists may be persuaded to adopt it. Often this claim cannot be legitimately made in the natural sciences. Moreover, it would be meaningless in Social Psychology since crises in that discipline occur for reasons that are fundamentally different from these in the natural sciences. A new paradigm in Social Psychology does not emerge as a solution to a crisis provoking problem, but rather it provokes a crisis because it solves a new problem, or because it solves an old problem in a new way. Consequently, it is the second claim that

appears to be central to the resolution of revolutions in both the natural sciences and Social Psychology. The new paradigm is likely to be adopted if it permits the prediction of phenomena that had been entirely unsuspected while the old one prevailed. In short, a paradigm may well succeed on grounds of novelty, simply by offering to replace an old and worn set of puzzles with a new set. For example, dissonance theory succeeded not because it solved problems posed by reinforcement theory, but rather because it was sufficiently unprecedented to attract an enduring group of adherents away from that and other competing modes of scientific activity. At the same time, it was sufficiently open-ended to leave all sorts of puzzles for the redefined group of practitioners to solve. If some form of self-perception theory replaces dissonance theory as a contemporary paradigm, this shift will have occurred for much the same reason. While it is clear that self-perception does not resolve all of the problems posed by dissonance theory, and a matter of debate as to whether it provides better solutions to the others, a decisive feature of the theory and its larger information processing/attributional framework is that it offers a unique and extensive set of puzzles for solution. These puzzles provide a basis for a new round of standard normal Social Psychology in which the paradigm can be articulated. To the extent that this activity is underway, it appears to be centering around conceptual and empirical distinctions between self and other's perceptions of the causes of behavior (for a review, see Jones and Nisbett, 1972). In addition, the new paradigm possesses sufficient intrinsic novelty to insure the continuance of a flashier brand of

puzzle-solving activity. Recent investigations of labelling errors in the development and treatment of emotional disorders provide examples (for reviews, see Winett, 1970, and Valins and Nisbett, 1972).

In conjunction with a new paradigm's capacity for solving puzzles, there is an important subjective consideration favoring the adoption of the self-perception approach. The self-perception theory has an aesthetic appeal which is lacking in the dissonance formulations. It is neat, parsimonious, and therefore elegant by comparison. Given that the decision to adopt a new paradigm must be based more upon future promise than past achievement, this aesthetic quality of the self-perception theory could well be a decisive factor in attracting social psychologists to the related set of puzzles. In a similar fashion, others may be attracted by Bem's debating skill. In Kuhn's view, the ultimate triumph of any new paradigm may depend upon such persons. If they fail to take up the paradigm for subjective reasons, it may never become sufficiently developed to attract the allegiance of a larger scientific community. Viewed in these terms, Bem's earlier observation on theoretical preference acquires a prescriptive significance beyond that which was originally intended. By noting that one's preference for theory has diminished to "a matter of loyalty or aesthetics", he not only signalled the hoped for end of a debate, but also identified a potentially important persuasive feature of his own theory.

It would be incorrect to conclude that Bem's work has precipitated the transfer of academic commitments from a motivational-consistency paradigm to an information processing/attributional paradigm. This

transition reflects a general trend in contemporary Psychology and has specific roots for Social Psychology in Schachter's (1964) early investigations of emotional labelling and Heider's (1958) common sense analyses of social perception. A conservative interpretation of the role played by the self-perception theory would be that of a highly visible historical marker for that transition in Social Psychology. A more generous, and I think more appropriate, interpretation is that the self-perception theory has had a crucial facilitatory or hastening effect upon this transition. Should the theory now be taken up and developed by even a few practitioners, it is likely that this latter interpretation will eventually be acknowledged by the redefined social psychological community. In the following chapter, I describe a set of three experiments which attempt to develop the self-perception theory as an initial step towards realizing this "prophecy".

CHAPTER THREE: EMPIRICAL STUDIES IN THE INFORMATION
PROCESSING/ATTRIBUTIONAL PARADIGM

So far I have argued that the recent history of attitudinal research can be viewed in terms of two conceptually distinct and sequentially separate periods of normal science and an intermediate period of revolutionary science. Festinger and Carlsmith's operationalization of the insufficient justification concept was suggested as a representative laboratory procedure for the earlier motivational-consistency normal research paradigm. In a related fashion, the simulation studies conducted by Bem and his critics constituted the representative laboratory procedure for research during the subsequent revolutionary period. While there is as yet no single laboratory procedure which is representative of the newer information processing/attributitional paradigm for normal research, Bem has developed a unique "credibility cueing" procedure for demonstrating the correctness of self-perception hypotheses. While it has been pretty much ignored in favor of simulations during the controversy, it is a useful procedure for doing research in the newer paradigm. What follows is an outline of this laboratory procedure and a description of three experiments conducted therein. The first experiment attempts to establish the internal validity of the cueing procedure and tests an hypothesis which is implicit in the self-perception theory. The second and third experiments attempt to broaden the generalizability of the cueing procedure and thereby indirectly explore its external validity. In a more general and important sense, these studies embrace both the theory

and procedure of self-perception in order to begin articulating the newer paradigm.

According to the self-perception theory, individuals come to know their own beliefs and other internal states partially by inferring them from observations of their own behavior and the circumstances under which it occurs. This general proposition is consistent with the findings of several studies in which individuals utilize available external cues in inferring such inner states as the emotions of happiness, anger (Schachter and Singer, 1962), and fear (Valins, 1966; Valins and Ray, 1967), as well as the perception of pain (Bandler, Madaras, and Bem, 1968; Zimbardo, Cohen, Weisenberg, Dworkin, and Firestone, 1969). In addition, the self-descriptive behavior of subjects in cognitive dissonance experiments which utilize monetary incentives is amenable to a cueing interpretation. Bem suggests that the money acts as either a "truth" or a "lie" signal to the subject. In the Festinger and Carlsmith experiment, for example, the \$20.00 incentive to tell a lie about the tasks cues the subject to attribute his verbal behavior ("I found these tasks fun and enjoyable.") to the incentive offered rather than to an internal predisposition such as a belief about the tasks. The large amount of money serves as a "lie" signal, telling the subject not to believe his statement. A subject who is offered only \$1.00 for the same verbal behavior is cued by the incentive to attribute his statement to an internal predisposition, or belief about the tasks. He does not attribute his statement to the incentive because it is of insufficient magnitude to warrant deceitful behavior. The small amount of money serves as a "truth" signal, telling the subject

to believe his statement. In order to explore this cueing notion, Bem (1966) has attempted to raise truth and lie cues "from birth" in the laboratory and to test their effects upon self-credibility.

Under the pretext of a study of lie detection, a subject completed a preliminary task about which he would later be required to tell some truths and some lies. A cover story informed the subject that the experimenter would tape record these statements and take various voice readings for the purpose of determining whether certain voice cues covary with successful detection of lies by a third person. The preliminary task involved crossing out 50 specified nouns from a list of 100 common nouns. The subject then underwent a training procedure in which he answered aloud a series of innocuous personal questions (e.g., "What is your name?"). After each question a tape recorder was turned on which automatically activated one of two colored lights in a ceiling fixture. The subject was instructed to answer the question truthfully whenever the amber light went on and to give a false answer whenever the green light went on. In this way he was presumably conditioned to believe himself whenever he spoke in the presence of the amber light and to not believe himself when he spoke in the presence of the green light. After this training session he was required to make statements about the preliminary task in the presence of the two lights. These statements were assigned by the experimenter so that instead of taking his cue from the lights, the subject simply verbalized a statement provided by the experimenter. Sometimes he was required to state that he had crossed out a word and sometimes to state that he had not crossed out

a different word (e.g., "I did not cross out the word TREE."). Unknown to the subject, half of the statements he was required to make were true and the other half were false. Again, the colored lights were connected to the tape recorder so that sometimes the amber light was on when he made his statement and sometimes the green light was on. Ostensibly, the lights were to serve the single purpose of indicating to the subject when he should make his statement by signalling when the tape recorder was on (i.e., when either light is on, the tape recorder is also on). In fact, however, it is intended that the lights will influence the subject's belief in the truthfulness of his statements. After each statement the subject indicated on a sheet of paper whether he recalled crossing out the word or not crossing it out. He also indicated how sure he was that his recall was accurate. According to the self-perception theory, the subject should tend to believe those statements that he makes in the presence of the light which had been previously paired with truth-telling, and to disbelieve those made in the presence of the light which had been previously paired with lie-telling. The pattern of recall data shown in Figure 3 would be consistent with these predictions. False statements in the presence of the "truth light" should result in more recall errors than false statements in the presence of the "lie light". Similarly, true statements in the presence of the "lie light" should result in more recall errors than true statements in the presence of the "truth light". The subject was also asked about a number of words that he had not made statements about. For these control words he was simply asked to recall whether or not he had crossed them out and how sure he

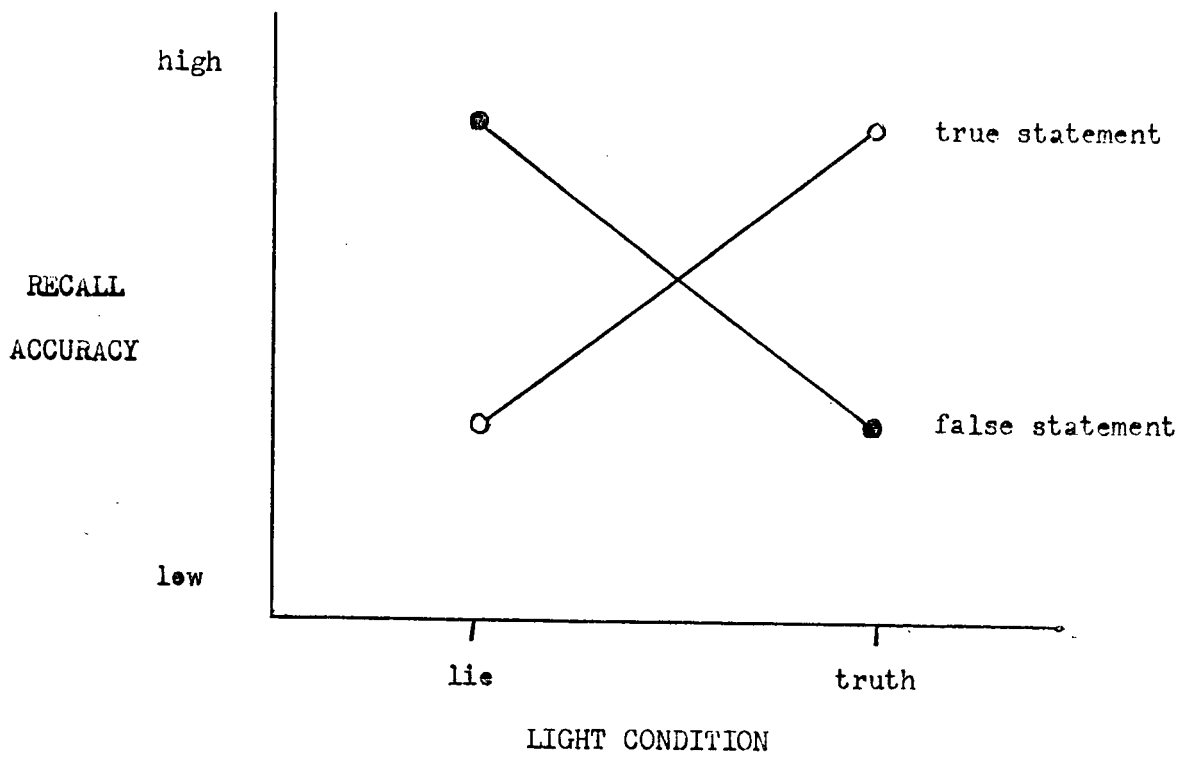


Figure 3. Recall accuracy predictions made by the self-perception theory.

was that his recall was accurate. No colored lights were activated in relation to these words. Presumably, recall of these words should show an intermediate number of errors.

To summarize, the credibility cueing procedure consists of three successive phases:

- Phase I - the subject engages in some activity which can later be used as the subject matter of his verbal statements.
- Phase II - the experimenter obtains personal information from the subject which he uses in establishing discriminative stimuli for truth and lie telling.
- Phase III - the discriminative stimuli are paired with statements which the subject is required to make concerning his previous activity in Phase I. The dependent measures are the subject's recall of what he actually did in Phase I and his confidence estimate for that recall.

These three phases can be termed task, training, and test phases, respectively.

The results of this study indicate that subjects' recall was influenced by their verbal behavior and the light cues in the directions predicted by the self-perception theory. Also, subjects' confidence estimates paralleled the recall data. When questioned, subjects were unable to verbalize the statement-light contingencies, and, in fact, most reported that they paid no attention to the lights during Phase III.

Although not tested directly by this procedure, the theory presumes that these results were mediated by differential self-credibility in the two light conditions.⁸

In each of the studies cited so far, the authors have assumed, apparently correctly, that the private internal cues associated with a particular inner state (an emotion, perception, attitude, or memory) are of a sufficiently indistinct nature that subjects would use external, public information in order to label it. The self-perception theory, however, appears to suggest a somewhat more precise relationship between the distinctiveness of internal cues and dependency upon external cues. Specifically, Bem contends that to the extent that information from internal cues is "weak, ambiguous, or uninterpretable", the individual must rely on external cues as signals of his own inner states. Although no formal postulate is advanced, this rationale implies that the two kinds of cues have a covariant relationship in the theory: when information from internal cues is indistinct, information from external cues will play a prominent role in the self-inference process; conversely,

⁸Bem (1965) has published one other study using the credibility cueing procedure. In that experiment he was able to influence subjective judgments of the "funniness" of cartoons in the manner predicted by his theory. Linder and Jones (1969) have conducted an experiment using this procedure and argued that the results can be accommodated by dissonance theory. This latter study is an example of the kind of "rear guard" action often undertaken by advocates of an earlier paradigm. It misses the point that the credibility cueing procedure and the related set of puzzles to which it can be applied were entirely unsuspected while the earlier paradigm prevailed. Although their demonstration may be of comfort to colleagues who are equally unlikely to experience the paradigm shift, it seems doubtful that it will have a significant effect upon the commitments of those who have embraced the newer paradigm.

when information from internal cues is distinct, information from external cues will play a less prominent role. The following experiment is designed to test this empirical prediction.

Experiment 1

A test of the proposed relationship between internal and external cues can be accomplished by introducing a third independent variable into the credibility cueing procedure. In addition to the two variables which Bem manipulated in his experiment, verbal statement cues (truth telling and lie telling) and light cues (truth signal and lie signal), the present design attempts to vary the distinctiveness of the internal cues available to subjects by requiring them to cross out two different kinds of stimulus items in the task phase of the experiment. In addition to crossing out words, as was done in Bem's experiment, subjects in the present study are required to cross out relatively meaningless trigrams as well. Because words are learned and recalled more readily than trigrams, it is assumed that any operations performed on words, such as crossing one out or not crossing it out, will be easier to remember than the same operations performed on trigrams. In self-perception terms, a task which requires an individual to either cross out or not cross out particular words should provide the individual with more distinctive internal cues concerning his own task behavior (presumably in the form of memory traces) than should a task which requires him to perform these operations on trigrams. If reliance upon external cues is an inverse function of the distinctiveness of internal cues, then recall of trigram operations should be more

susceptible to systematic distortion by verbal statements emitted in the presence of discriminative light stimuli than should recall of word operations. According to the self-perception theory, the predicted recall distortions will have the following two cueing components:

(a) false statements made in the presence of the truth light will result in significantly poorer recall than false statements made in the presence of the lie light, and (b) true statements made in the presence of the lie light will result in significantly poorer recall than true statements made in the presence of the truth light. The major prediction is that these distortions will be more pronounced for recall of stimulus operations designed to produce indistinct internal cues (crossing out trigrams) than for recall of stimulus operations designed to produce distinct internal cues (crossing out words).

In a recent experiment, Maslach (1971) failed to replicate Bem's findings. Instead of the interaction which Bem obtained, she observed that her subjects' recall was more accurate in both lie light conditions independent of the truthfulness of their statements. In order to account for this finding she speculated that the lie light may have acquired the properties of a "vigilance" cue by virtue of the relatively complex reasoning it demands in the training phase of the experiment (the subject must first think of the correct answer to a question, and then generate one of many possible incorrect alternatives). In the subsequent recall phase, this light could cue the subject to take his time and think carefully about his response. The effect of this "vigilance" would be fewer recall errors in the lie light condition. In order to rule out

this possible methodological problem, in the training phase of the present study subjects are required to think of a false as well as a true answer to each question prior to the activation of either light. Of course, they are only required to give the false answer when the lie light is illuminated.

A second difficulty, the possibility that Bem's results were mediated by a process other than the one he postulates, can be minimized by subjecting the results of the present experiment to a signal detection analysis. Although Bem assumes that the pattern of recall errors he obtained was due to differential credibility cueing in the various conditions of his experiment, it is possible that these data merely reflect differential guessing rates. According to this view, when a subject makes a false statement in the presence of the truth light, or a true statement in the presence of the lie light, he may become confused by the "contradiction" between the information provided by the light cue and any information that he possesses in the form of memory traces of his previous behavior. This source of confusion, which would be absent in the false statement-lie light and true statement-truth light conditions, may predispose the subject to increase his rate of guessing. The effect of such a differential guessing rate would be a closer approximation of chance responding in the false statement-truth light and true statement-lie light conditions, the two conditions in which Bem finds the poorest recall. If differential guessing rates did in fact occur in Bem's experiment, the effect would be confounded with any self-perception effects by virtue of the type of data used as a

recall measure in that experiment. That measure combined two categories of recall errors: items which the subject originally crossed out and subsequently recalled as not crossed out (misses), plus items which the subject originally did not cross out and subsequently recalled as crossed out (false alarms). In order to correct recall data for differential guessing rates it is necessary to separate these two response categories and employ a measure which takes into account the remaining possible response categories: correct recall of items which the subject did not cross out (correct rejections) and correct recall of items which the subject did cross out (hits). One such measure commonly used in psychophysics and human memory research is d' (Swets, 1964; Kintsch, 1970). In signal detection terms, d' is an index of an observer's sensory capacities, or of effective signal strength. In the present study d' is employed as a measure of how accurately a subject reconstructs past events when a guessing factor is removed.

Method

Subjects were four male and seven female undergraduates. After being seated at a table containing a panel with a microphone between two colored light bulbs (red and green), a desk lamp, and a large opaque screen, the subject was given a two-page list of stimulus items which contained a random sequence of 100 common nouns interspersed with 100 low-meaningfulness CVC trigrams (from Archer's (1960) re-evaluation of meaningfulness of all possible CVC trigrams). The subject was also given two alphabetical lists, one containing 50 of the nouns and the

other containing 50 of the trigrams, and instructed to draw a line through each word and each trigram on the long list that also appeared in the alphabetical guides. After completing this preliminary task, the subject filled out a 50-item information form which asked innocuous personal questions such as, "What is your favorite sport?", "Did you watch television last night?", etc. After obtaining the completed questionnaire, the experimenter positioned himself across the table from the subject and behind the opaque screen. A training procedure was then employed to establish the two colored lights as discriminative stimuli for truth telling and lie telling. A lie-detection cover story informed the subject that his voice would be recorded and various readings taken in order to determine whether lies could be detected in this way. The experimenter then asked questions one at a time from the information form, and a few seconds after each question one of the two colored lights was illuminated. The subject was required to answer the question truthfully whenever the red light went on and untruthfully whenever the green light went on. For some subjects the two lights were reversed so that the green light signalled a true response and the red light signalled an untrue response. In order to rule out a possible "vigilance artifact", subjects were instructed to think of an untrue answer during the pause between the question and the onset of a light. Half of the questions required true responses and half required untrue responses.

The final phase of the study tested the effects of this training procedure on recall performance. The subject was required to make statements about the words and trigrams that he did or did not cross out

in the first part of the study. Fifty words and 50 trigrams from the 200-item list were employed in this phase of the experiment. Using a predetermined schedule, the experimenter announced a word or a trigram and verbally cued the subject to either state that he had or had not crossed out that stimulus item earlier. The subject was instructed to silently rehearse the statement he was about to make until the recording equipment was turned on, and then give the statement aloud (e.g., "I did cross out the item TREE."). So that he would know when the equipment was turned on, the experimenter indicated that he would leave the colored lights hooked up to the equipment and these would continue to flash on and off in a programmed random sequence. The rehearsal procedure was introduced in order to prevent the subject from attempting to recall the original stimulus operation prior to his making a verbal statement in the presence of a light cue. On half of the trials there was a two-second delay between the verbal cue from the experimenter and the illumination of a light, and on the other half of the trials this delay was four seconds. The reason for using two intervals was to maximize the cue value of the colored light. It was felt that a constant delay between the experimenter's cue and the light cue might predispose subjects to use the time interval rather than the light as a signal to make their verbal statements. When the subject completed his verbal statement the colored light was turned off and the desk light turned back on. The subject then verbally indicated to the experimenter whether he recalled crossing out that stimulus item or not crossing it out earlier. He also indicated how confident he was in the accuracy of

his recall on a five-point scale. In addition, on some of the trials the experimenter verbally cued the subject to simply recall whether he had crossed out a stimulus item or had not crossed it out earlier, and to give a confidence estimate. On these trials the subject did not make a verbal statement and no colored light was activated.

A total of 20 stimulus items, 10 words and 10 trigrams, were employed in each of the following conditions: true statement-truth light, true statement-lie light, false statement-truth light, false statement-lie light, and control (no statement-no light). Half of the stimulus items in each condition had actually been crossed out by the subject; half had not been crossed out. The two control conditions were not varied factorially with the statement cue and light cue variables. Consequently, data from the control conditions is excluded from the overall analyses and employed only in subsequent comparisons. This procedure yields a complete $2 \times 2 \times 2$ within-subjects design with 11 observations per cell on each of two dependent measures, recall accuracy and confidence in recall accuracy.

Results and Discussion

Table 1 presents the mean proportions of hits (the number of crossed out items correctly recalled as crossed out, divided by the total number of crossed out items presented in the test phase) and false alarms (the number of not crossed out items incorrectly recalled as crossed out, divided by the total number of not crossed out items presented in the test phase), as well as the corresponding d' values

for each of the eight experimental conditions. The higher the d' score, the more accurate is subjects' recall. For example, in the true statement-truth light condition for words, where subjects are highly accurate, the mean d' score is +1.77; in the false statement-truth light condition for trigrams, where recall accuracy is poor, the mean d' score is -0.48. The negative sign on the latter d' score denotes a false alarm rate which exceeds the hit rate in that condition. The mean confidence ratings for each of the experimental conditions are shown in Table 2. Separate analyses of variance were performed on the d' data and on the confidence data. These analyses are summarized in Tables 3 and 4. They indicate a main effect for stimuli on both dependent measures ($F = 6.90$, df 1/10, $p < .05$ for recall data; $F = 235.30$, df 1/10, $p < .001$ for confidence data). This suggests that the manipulation of internal cue distinctiveness was successful. In the experimental conditions, subjects' recall performance was poorer and they were less confident in the accuracy of their recall for trigram operations ($\bar{X} = 0.45$ for d' and $\bar{X} = 2.76$ for confidence) than for word operations ($\bar{X} = 1.69$ for d' and $\bar{X} = 3.20$ for confidence). Although recall of control words is relatively poor in comparison with recall of experimental words, the control results parallel those observed in the experimental conditions. In the absence of statement and light cues, subjects' recall performance yielded mean d' values of 0.52 and 0.97 in trigram and word conditions respectively. This difference approached conventional levels of significance ($t = 1.52$, df 10, $p < .07$).⁹

⁹ All t tests are one-tailed.

Table 1. Mean proportions of hits, false alarms, and corresponding d' values for recall of word and trigram operations in each statement and light cue condition.

Statement Condition		Stimulus Condition			
		WORDS		TRIGRAMS	
		Light Condition		Light Condition	
		TRUTH	LIE	TRUTH	LIE
	Hits	.85	.79	.66	.45
TRUE	False Alarms	.43	.31	.40	.36
	d'	1.77	1.77	1.07	0.16
	Hits	.76	.68	.35	.66
FALSE	False Alarms	.27	.35	.51	.40
	d'	1.79	1.44	-0.48	1.05

Table 2. Mean ratings of confidence in recall accuracy for word and trigram operations in each statement and light cue condition.

Statement Condition		Stimulus Condition			
		WORDS		TRIGRAMS	
		Light Condition		Light Condition	
		TRUTH	LIE	TRUTH	LIE
	TRUE	3.77	3.92	2.84	2.90
	FALSE	3.74	3.76	2.56	2.75

Table 3. Analysis of variance of recall scores (d').

Source	df	MS	F
Total	87		
Subjects	10		
Stimulus (A)	1	33.87	6.90*
Statement (B)	1	1.29	2.43
Light (C)	1	0.10	0.10
A x B x C	1	18.11	12.84**
A x B	1	0.16	0.13
A x C	1	1.33	0.88
B x C	1	6.02	6.62*
error A	10	4.91	
error B	10	0.53	
error C	10	1.05	
error A x B x C	10	1.41	
error A x B	10	1.25	
error A x C	10	1.52	
error B x C	10	0.91	

* $p < .05$ ** $p < .005$

Table 4. Analysis of variance of confidence scores.

Source	df	MS	F
Total	87		
Subjects	10		
Stimulus (A)	1	23.53	235.30**
Statement (B)	1	0.51	5.10*
Light (C)	1	0.25	2.27
A x B x C	1	0.16	0.73
A x B	1	0.07	0.77
A x C	1	0.01	0.13
B x C	1	0.01	0.25
error A	10	0.10	
error B	10	0.10	
error C	10	0.11	
error A x B x C	10	0.22	
error A x B	10	0.09	
error A x C	10	0.08	
error B x C	10	0.04	

* $p < .05$ ** $p < .001$

The mean confidence ratings were 2.90 in the control trigram condition and 3.84 in the control word condition. This difference was statistically significant ($t = 5.28$, df 10, $p < .0005$). In addition to confirming the success of the distinctiveness manipulation, the results of these analyses showed no main effect for light cues on either dependent measure. Since Maslach's "vigilance" hypothesis would predict superior recall performance in the lie light conditions (and possibly greater confidence in recall accuracy), the absence of any effects for lights suggests that either this "artifact" may not exist, or that it was removed by the procedural safeguard employed in the present study.

We can now turn to the self-perception predictions involving recall accuracy. The analysis of variance on the d' scores revealed an interaction effect of statement and light cues ($F = 6.62$, df 1/10, $p < .05$) as well as an interaction effect of stimulus, statement, and light cues ($F = 12.84$, df 1/10, $p < .005$). The significant two-way interaction is consistent with the general prediction of poorer recall performance in these conditions where the light provides misleading information about the validity of a subject's statements. The mean d' value for the combined false statement-truth light and true statement-lie light conditions is 0.81 as compared with a mean d' value of 1.33 in the combined true statement-truth light and false statement-lie light conditions. Similarly, the significant three-way interaction is consistent with the major prediction of a more pronounced effect of external cues on recall of trigram operations. However, an inspection of these data, shown in Figure 4, suggests that these two interaction effects provide

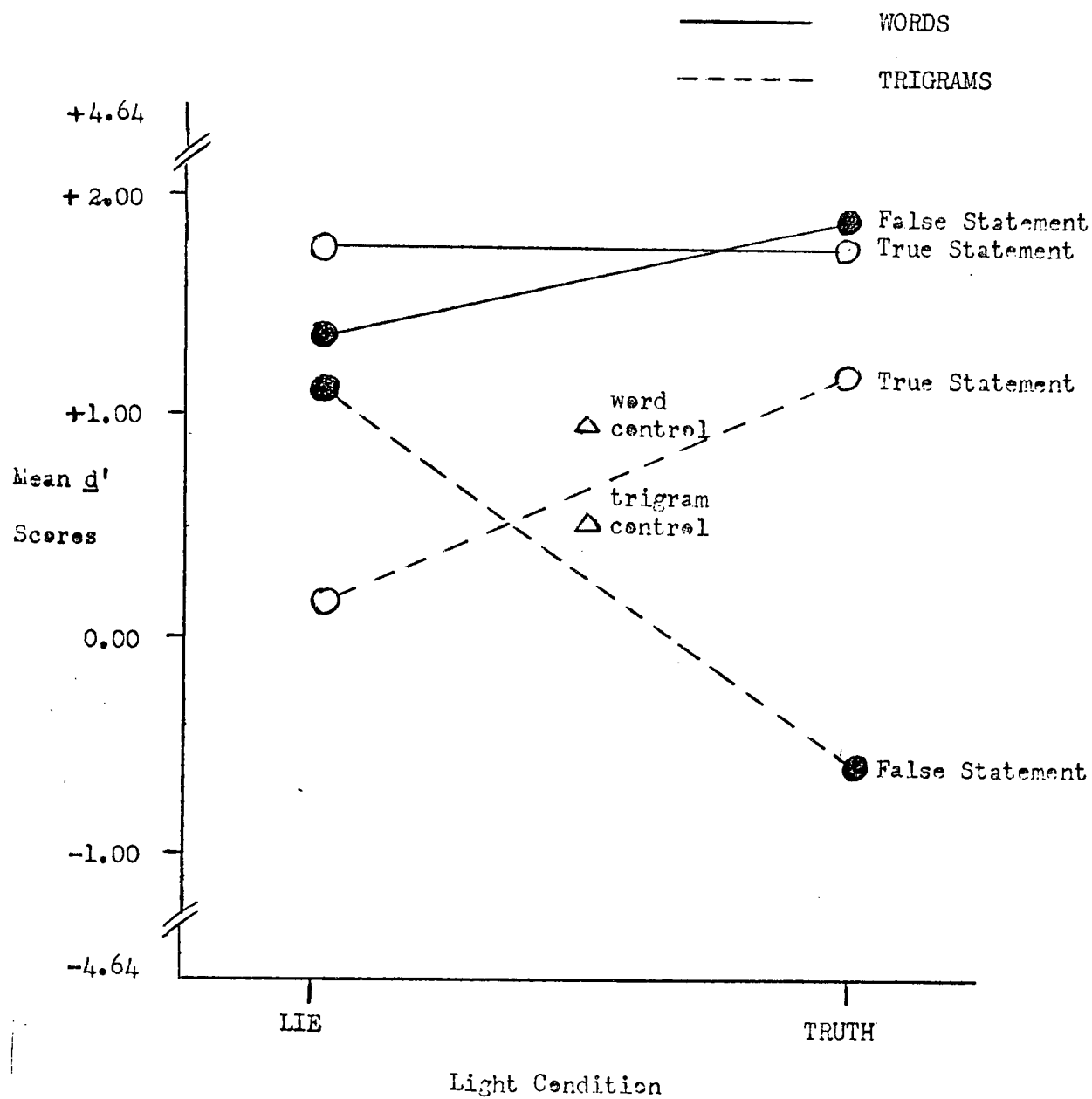


Figure 4. Recall accuracy for word and trigram operations in each statement and light cue condition.

only qualified support for the predictions. The significant statement x light interaction appears to be entirely due to the trigram stimulus variable, and the significant stimulus x statement x light interaction appears to reflect the presence of the trigram effect and the absence of a word effect, rather than the presence of a strong trigram effect and a relatively weaker word effect as predicted in the major hypothesis. In order to further explore these possibilities, d' scores from the word and trigram experimental conditions were analyzed separately using orthogonal linear contrasts. The results of the analysis of the word data confirmed the absence of a statement x light interaction ($F < 1.00$, df 1/10). In addition, there were no main effects of these two variables ($F < 1.00$, df 1/10 for statements; $F < 1.00$, df 1/10 for lights). The results of the analysis of the trigram data revealed the expected interaction between statement and light cues ($F = 11.56$, df 1/10, $p < .01$) and absence of main effects ($F < 1.00$, df 1/10 for statements; $F = 1.19$, df 1/10 for lights). This confirms the notion that errors in the recall of trigram operations but not in the recall of word operations produced the significant interaction effects. The results of a further analysis of variance on the word data using uncorrected recall scores (hits and correct rejections combined) also failed to reveal a significant interaction. This analysis is summarized in Table 5. Taken together, the results of these analyses suggest that the failure to obtain a significant interaction for recall of word operations was not simply due to the adoption of a guessing-corrected recall measure in the present experiment. It follows that this failure to "replicate" Bem's cueing

Table 5. Analysis of variance of uncorrected recall
scores for words.

Source	df	MS	F
Total	43		
Subjects	10		
Statement (B)	1	0.12	0.03
Light (C)	1	1.08	1.06
B x C	1	2.90	1.30
error B	10	0.42	
error C	10	1.02	
error B x C	10	2.23	

effect for words cannot be taken as presumptive evidence for a "differential guessing" interpretation of his original findings. This conclusion is underscored by the significant interaction obtained with guessing-corrected scores in the trigram condition of the present experiment. If anything, the present demonstration of a cueing effect with the possibility of "vigilance" and differential guessing artifacts removed would seem to establish the internal validity of the credibility cueing procedure more firmly than was previously done by Bem.

Since Bem's interaction was not replicated for recall of word operations, the results of the present experiment cannot possibly support a strict interpretation of the major prediction. Although subjects' recall of trigram operations reflected a significant effect of external cues, the absence of any such effect in the recall of word operations renders meaningless any comparisons of relative magnitude of effects. The major prediction receives qualified support, however, since external cues had their effect in the weak internal cue condition and not in the relatively stronger internal cue condition. These are the conditions which should be most likely and least likely respectively to produce effects according to the present interpretation of the self-perception theory.

Because external cues had an effect on recall performance in the trigram condition only, individual comparisons between trigram and word conditions were not subjected to statistical test. Instead, only the trigram data were analyzed further in order to examine the hypothesized components of the external cue effect. Data from the control condition

were also considered in these analyses. Tables 6 and 7 compare the effects of false statements and true statements separately. The results of these comparisons provide support for both of the hypothesized components. False statements made in the presence of the truth light resulted in significantly poorer recall performance than did false statements made in the presence of the lie light ($t = 2.21$, $df\ 10$, $p < .05$). Similarly, true statements made in the presence of the lie light resulted in significantly poorer recall performance than did true statements made in the presence of the truth light ($t = 3.12$, $df\ 10$, $p < .01$). Although differences between control and experimental conditions were all in the expected directions, only that difference between the control and the false statement-truth light condition achieved the conventional .05 level of significance ($t = 2.10$, $df\ 10$). The self-perception interpretation of these results is that cues previously associated with truth telling induce a person to believe his own verbal statements made in their presence. Hence, in the trigram truth light condition, believing false statements leads subjects to be less accurate, while believing true statements leads them to be more accurate. Similarly, the theory holds that cues previously associated with lie telling induce a person to disbelieve his own verbal statements made in their presence. Hence, in the trigram lie light condition, disbelieving false statements leads subjects to be more accurate, while disbelieving true statements leads them to be less accurate.

Finally, the analysis of variance of the confidence data revealed a main effect for statements ($F = 5.10$, $df\ 1/10$, $p < .05$) in addition to

Table 6. Recall accuracy comparisons for trigram operations following false statements.

Light Condition	Mean d' Value
Truth (A)	-0.48
Lie (B)	1.03
Control (C)	0.52
	t
A vs. B	2.21*
A vs. C	2.10*
B vs. C	0.94

* $p < .05$

Table 7. Recall accuracy comparisons for trigram operations following true statements.

Light Condition	Mean d' Value
Truth (A)	1.07
Lie (B)	0.15
Control (C)	0.52
	t
A vs. B	3.12**
A vs. C	1.21
B vs. C	0.86

** $p < .01$

the stimulus effect discussed earlier. Subjects were more confident in the accuracy of their recall after making true statements ($\bar{X}=3.76$) than after making false statements ($\bar{X}=3.20$), an intuitively plausible finding. When confidence estimates were examined in the various statement and light conditions within each stimulus condition using Dunn's test (Kirk, 1970), no significant differences were found. In addition, there were no significant differences between any of these experimental conditions and the control conditions. Although these findings are inconsistent with Bem's observation that confidence estimates parallel recall scores (i.e., those conditions which produced the most recall errors in his experiment also produced the lowest confidence estimates), it is not necessarily inconsistent with his theory. In fact, if subjects actually believed their false statements in one condition and disbelieved their true statements in the other, then confidence estimates should not differ across these conditions.

When questioned at the completion of the experiment seven subjects indicated that the statements and lights had no effect upon their recall and the remaining four subjects indicated that the lights may have confused them. None of the subjects was able to verbalize a systematic relationship between the lights and the truth value of the statements.

It is unclear why Bem's findings were not replicated in the word condition of the present study. One possible explanation involves the rehearsal procedure used to prevent subjects from attempting to recall prior to making statements in the presence of a light. This procedure may have been ineffective in delaying recall of operations for stimuli

which produced distinctive memory cues. Because the subjects possessed relatively distinct internal cues for word operations, recall may have been more or less automatic once the experimenter verbalized the cue word. Consequently, subjects would be predisposed to disregard subsequent information from their statement and the light. However, the rehearsal procedure may have been more effective in delaying recall of trigram operations since these operations produced relatively indistinct internal cues which, by definition, would not be as readily accessible as the more distinctive word cues. In short, recall of these stimulus operations which produce highly distinctive internal cues may have capitalized on the recognition aspect of the present task, and thereby eliminated, or at least reduced, any need for additional input from external cues. This explanation would account for the presence of an external cue effect for recall of trigram operations and the absence of an effect for word operations. Of course, it does not explain how Bem obtained an external cue effect for recall of word operations in his experiment.

Experiment 2

Given that subjects partially infer their own inner states from external cues which have been designed and manipulated by an experimenter, one might speculate as to the sort of naturally occurring external cues which could serve this function for people in the real world. An obvious possibility for a social psychologist to consider is that the presence of other people may provide an individual with cues to his inner states.

For example, depending upon an individual's history of interaction with particular other persons, these others may acquire "truth" or "lie" cue properties for him in much the same way as has been demonstrated for colored lights. A partial approximation of a real world social situation in which people might serve this credibility cue function is attempted in this experiment.

The design is similar to that of the first experiment with two major exceptions. Videotaped interpersonal cues are employed in place of the colored light cues used in the first experiment. These cues consist of two different interviewers who appear full-face on a television monitor and pose questions in an "interviewing" format. In addition, the stimuli are not manipulated as an independent variable in this experiment. Instead, trigram stimuli are used exclusively since a cueing effect was previously obtained for these stimuli, but not for word stimuli. The self-perception theory predicts the same pattern of recall distortions that was obtained for trigram stimuli in the first experiment, an interaction of the two independent variables, statement cues and interviewer cues. False statements made in the presence of the truth interviewer should result in significantly poorer recall than false statements made in the presence of the lie interviewer, and true statements made in the presence of the lie interviewer should result in significantly poorer recall than true statements made in the presence of the truth interviewer.

Method

Subjects were 11 male undergraduates. After being seated at a table containing a large television monitor, television camera, microphone, and desk lamp, the subject completed a crossing out task (50 of the same 100 trigrams used in Experiment 1) and filled out the 50-item information questionnaire according to the procedure described in Experiment 1. After obtaining the completed questionnaire, the experimenter instituted a training procedure to establish two videotaped interviewers as discriminative stimuli for truth telling and lie telling. A lie-detection cover story informed the subject that he would be recorded on videotape in order to determine whether observers could detect lies by observing and listening to the tape. In order to create an impression that he was being videotaped, the experimenter focussed the camera on the subject and fed this input into the television monitor which faced him. This input to the monitor was then replaced by a short videotape of two different interviewers, each asking three innocuous questions (e.g., "Do you like coffee?"). In the tape an interviewer appeared full-face, asked a question, remained on the monitor just long enough for an answer to be given, and then disappeared from the monitor. After a period of a few seconds, he then either reappeared, or a different interviewer appeared, and the process was repeated with a different question. The subject was asked to watch this 6-question tape without responding. Upon completion of the tape he was requested to choose the interviewer that he would answer truthfully when subsequently asked questions from the 50-item questionnaire. The interviewer not chosen was, of course,

to be answered untruthfully. If the subject indicated that he had no preference, an assignment was made by the experimenter. The 6-question tape was then replayed and the subject responded to each question either truthfully or untruthfully according to his choice (or assignment). If the subject made no errors in this practice session, the 50-question tape was then played. This tape consisted of each interviewer asking 25 questions with the order of the interviewers' appearance randomized. Hence, half of the questions required true responses and half required untrue responses. The subject was instructed to look right at the interviewer on the monitor when answering a question. As in the first experiment, he was instructed to have an untrue answer ready to give for each question. To facilitate this request the subject was given a list of the 50 questions in the order in which they appeared on the tape. In this way he could anticipate each question before it was asked and prepare an untrue answer. Of course, he only gave this answer when the lie interviewer asked the question.

The final phase of the study tested the effects of this training procedure on recall performance. The subject was required to make statements about the trigrams that he did or did not cross out in the task phase of the experiment. Forty trigrams from the 100-item list were employed in this phase of the experiment. On each of the 40 trials an interviewer appeared on the monitor, held up a 3" x 5" card with a trigram printed on it, and asked "Did you cross out this item?". Within a few seconds he disappeared from the screen and five seconds later he either reappeared, or the other interviewer appeared, and this

procedure was repeated with a different trigram. A number between 1 and 100 was printed in the top right-hand corner of each card. The subject was instructed to answer "Yes, I did cross out the item (spells out the trigram)." whenever the number was less than 50, and to answer "No I did not cross out the item (spells out the trigram)." whenever the number was greater than 50. The code was reversed for the last 20 trials. No rehearsal procedure was employed in this experiment. Instead, the subject was instructed to make his response as soon as possible after the interviewer finished asking the question. The subject then verbally indicated to the experimenter whether he recalled crossing out that stimulus item or not crossing it out earlier. Confidence estimates were not obtained in this experiment. In order to maintain the interpersonal interactional property of the experiment, control items were not interspersed with experimental items in the test phase. Instead, after completion of the 40 experimental trials the experimenter sequentially showed the subject ten trigrams and asked him to recall whether or not he crossed each one out.

A total of ten stimulus items were employed in each of the following conditions: true statement-truth interviewer, true statement-lie interviewer, false statement-truth interviewer, false statement-lie interviewer, and control (no statement-no interviewer). Half of the stimulus items in each condition had actually been crossed out; half had not been crossed out. Data from the control condition is excluded from the overall analysis and employed only in subsequent comparisons. This procedure yields a complete 2 x 2 within-subjects design with 11 observations per

cell on the dependent measure, recall accuracy.

Results and Discussion

Table 8 presents the mean proportions of hits and false alarms, as well as the corresponding d' values for each of the four experimental conditions. An analysis of variance performed on the d' data is summarized in Table 9. It indicates a main effect for interviewers which is highly significant ($F = 13.14$, df 1/10, $p < .005$), a main effect for statements which approaches conventional significance levels ($F = 4.59$, df 1/10, $p < .07$), and an interaction between these two variables which approaches conventional significance levels ($F = 3.26$, df 1/10, $p < .11$). An inspection of these data, shown in Figure 5, suggests that these effects are due to the relatively high level of recall accuracy in the true statement-truth interviewer condition ($\bar{X} = 1.45$ as compared with \bar{X} s of 0.41, 0.17, and 0.13 in the other conditions). Because the interaction approached conventional levels of significance, a Newman-Keuls test was employed on the experimental means in order to confirm this observation. Recall accuracy was significantly higher in the true statement-truth interviewer condition than in any of the other three conditions ($p < .01$). In addition, there were no significant differences between the other three conditions. A comparison of the control mean ($\bar{X} = 1.13$) with the experimental means using Dunnett's test revealed that recall accuracy was significantly higher in the absence of statement and interviewer cues than it was in the presence of such cues for three of the four comparisons ($p < .05$). The single exception was, of course,

Table 8. Mean proportions of hits, false alarms, and corresponding d' values for recall in each statement and interviewer cue condition.

Statement Condition		Interviewer Condition	
		TRUTH	LIE
TRUE	Hits	.76	.45
	False Alarms	.42	.42
	d'	1.45	0.17
FALSE	Hits	.47	.49
	False Alarms	.34	.47
	d'	0.41	0.13

Table 9. Analysis of variance of recall scores (d').

Source	df	MS	F
Total	43		
Subjects	10		
Statement (A)	1	3.21	4.59
Interviewer (B)	1	6.70	13.14*
A x B	1	2.74	3.26
error A	10	0.70	
error B	10	0.51	
error A x B	10	0.84	

* $p < .005$

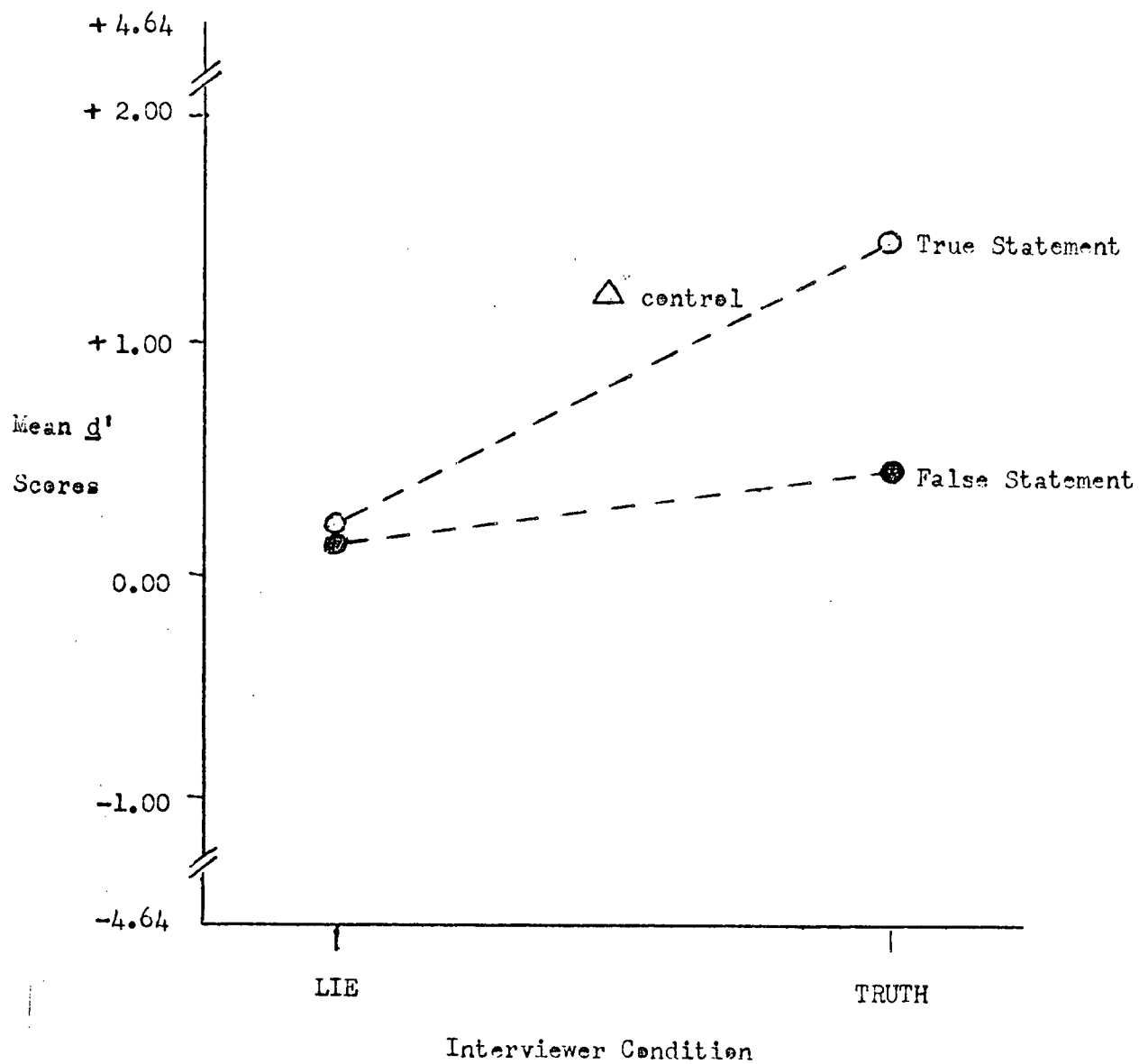


Figure 5. Recall accuracy in each statement and interviewer cue condition.

the comparison between the means of the control and true statement-truth interviewer conditions.

There are two possible credibility cueing interpretations of these data: (a) an interviewer cueing effect was produced for true statements (recall accuracy higher for truth interviewer than for lie interviewer) but not for false statements, or (b) a statement cueing effect was produced for the truth interviewer (recall accuracy higher for true statements than for false statements) but not for the lie interviewer. Although the present experiment provides no direct evidence to support a choice between these two alternative interpretations, the latter is favored for presumptive reasons. Recall that the trigram results of Experiment 1 supported both of the hypothesized components of the cueing effect (behavior and situational cues). Since the present experiment retains the behavioral cues (verbal statements) employed in the first experiment, but replaces impersonal situational cues (colored lights) with interpersonal situational cues (videotaped interviewers), this altered situational cueing variable would be the most likely source of differences between the results of the two studies. In this connection, it may be that interpersonal stimuli, in particular the lie interviewer of the present study, either do not acquire credibility cueing properties as readily as do impersonal stimuli, or if they do readily acquire such properties their effects are fragile. With respect to the proposed absence of a statement cueing effect for the lie interviewer, it may be that telling a series of lies to another person, even in the present mechanized situation, results in something more (or other) than the

attachment of (in)credibility cues to that stimulus person. This is an unusual and possibly anxiety-provoking bit of verbal behavior, the effects of which might become conditioned to the stimulus person's presence. This anxiety could interfere with the acquisition of credibility cueing properties by the stimulus person, and, once conditioned, override any possible cueing effects. This would result in a uniformly (i.e., independent of statement cue) low level of recall accuracy in the presence of the stimulus person. The observation that recall accuracy was significantly higher in the control condition is consistent with this speculation.

To summarize this tentative interpretation of the present results, cues previously associated with truth telling induced subjects to believe their own verbal statements made in the presence of these cues, but other interpersonal cues previously associated with lie telling failed to induce subjects to disbelieve their own verbal statements made in the presence of these cues. Unfortunately, this interpretation is further complicated by a procedural problem in the present experiment. The procedure employed here allowed subjects to choose the interviewer to whom they would give truthful responses. Of the 11 subjects, eight chose to answer interviewer A truthfully, one chose to answer interviewer B truthfully, and the remaining two were assigned to answer B truthfully when they indicated no preference. Consequently, one interviewer (A) served as the truth cue for three-quarters of the subjects, and the other interviewer (B) served as the lie cue for these same subjects. The interviewers' roles were reversed for the remaining subjects. This

procedure introduces a confound between the credibility cueing effects which the experimenter is attempting to produce and any "interpersonal" cueing effects that a particular interviewer might produce by virtue of his own appearance, speech, manner, etc.¹⁰ Because there is some presumptive evidence for an interpersonal cueing effect in the present experiment, this is a real problem. While neither lie interviewer produced a statement cueing effect (with lie interviewer A, one of three subjects showed the effect; with lie interviewer B, two of eight subjects showed the effect), it appears that the significant statement cueing effect was produced entirely by one of the two truth interviewers. With truth interviewer A, eight of eight subjects showed the effect, whereas only one of three subjects showed the effect with truth interviewer B. Although the small sample precludes the drawing of any firm conclusions, these observations raise the possibility that some stimulus persons are more likely to produce credibility cueing effects than others, presumably by virtue of some unspecified interpersonal characteristics.

Experiment 3

Although deficiencies in the design of Experiment 2 severely limit the value of the results, an intriguing speculation was advanced to the

¹⁰ Since two of three subjects were assigned to answer truthfully to interviewer B, and none were assigned to interviewer A, the effects of "choice" would also be confounded with any credibility cueing effect.

effect that different stimulus persons could have different capacities for producing credibility cueing effects. A third experiment attempts to capitalize upon this serendipitous effect of the experimenter's prior sleepiness by examining this possibility more closely. In addition to employing interviewers as credibility cues as was done previously, the interviewer's identity is manipulated as an independent variable in this experiment. Second, in an attempt to increase both laboratory and mundane realism, live interviewers replace the relatively mechanical and impersonal television images of the previous experiment. Finally, the internal cue distinctiveness hypothesis is retested in this experimental setting.

The design and procedure are similar to those of the first experiment with the following exceptions. Live interviewers are employed as truth and lie cues in place of colored lights, and this independent variable is manipulated between subjects instead of within subjects. Two subjects are run in an experimental session instead of one. After independently completing the crossing out task, they are both conditioned to believe statements which they make to one interviewer and to disbelieve statements which they make to the other interviewer. In the test phase of the experiment the two subjects are separated so that one is tested in the presence of the truth interviewer and the other is tested in the presence of the lie interviewer. The interviewer's identity is manipulated as an independent variable between subjects by having each of the two interviewers serve as a truth cue for one-quarter of the subjects and as a lie cue for one-quarter of the subjects. This manipulation provides

a weak experimental examination of the question raised by Experiment 2 concerning the differential capacity that various stimulus persons might have for producing credibility cueing effects. If the two interviewers employed in the present study have different capacities for producing credibility cueing effects, this could be manifested in a number of different ways, including the presence of an effect for one interviewer and the absence of an effect for the other, the presence of effects of different magnitudes, partial effects, etc. Of course, if there are no differences between the effects produced by the two interviewers, this would only indicate that differential cueing effects are not produced by these two stimulus persons. The question of whether or not other stimulus persons could produce differential cueing effects would still be unresolved.

In order to retest the hypothesis relating to internal cue distinctiveness which received partial support in Experiment 1, stimuli are manipulated as an independent variable in the present experiment. Again, it is predicted that recall of trigram operations will be more susceptible to systematic distortion by verbal statements emitted in the presence of discriminative interviewer stimuli than will recall of word operations.

Method

Subjects were 40 male undergraduates who participated in pairs in each experimental session. They were seated across from each other at a table in the center of a 7' x 18' room. An opaque screen separated the two subjects. They independently completed a crossing out task

(50 of the same 100 words and 50 of the same 100 trigrams employed in Experiment 1) and filled out a 48-item information questionnaire according to the procedure described in Experiment 1. After obtaining the completed questionnaires, the experimenter instituted a training procedure to establish two live interviewers as discriminative stimuli for truth telling and lie telling. A lie-detection cover story informed subjects that they would be asked the series of questions from the information form by the two interviewers who would attempt to distinguish between truthful and untruthful answers by attending to the subjects' voices, facial expressions, etc. Both subjects were instructed to answer the same interviewer truthfully and to answer the other interviewer untruthfully. These roles were assigned according to a prearranged schedule so that 20 subjects (10 pairs) answered interviewer A truthfully and interviewer B untruthfully, and 20 subjects (10 pairs) answered B truthfully and A untruthfully. Subjects were instructed to look right at an interviewer when answering his questions and to respond with a complete sentence, ostensibly for the purpose of providing an interviewer with a good sample of behavior from which he could make a judgment concerning truthfulness.

After allowing subjects a few minutes to reread the 48 questions from the information form and to practice preparing untrue answers, the two interviewers were brought into the room and introduced to the subjects. One interviewer was seated behind a desk at one end of the room and the other interviewer was seated behind a desk at the opposite end of the room. A subject could observe one interviewer by turning his head to

the left, and the other by turning his head to the right. Each interviewer asked half of the questions from the information form according to the following format: an interviewer addressed a question first to one subject, received an answer, scribbled on a note pad to create the impression that he was making a judgment about the truthfulness of the answer he had just received, and then repeated the process by addressing the same question to the other subject. The order of the interviewers' questioning was randomized, as was the selection of the subject who was to answer a question first.

The final phase of the study tested the effects of this procedure on recall performance. The two subjects were separated and each one was tested independently in a small room by one of the two interviewers. Ten subjects were tested in each of the four conditions created by the between-subjects manipulations of the interviewer cue and interviewer identity independent variables: truth interviewer A, lie interviewer B, truth interviewer B, and lie interviewer A. Statements and stimuli were manipulated as independent variables within subjects, so that each subject was required to make true and false statements about the words and trigrams that he did or did not cross out in the task phase of the experiment. Twenty words and 20 trigrams were employed and half of the statements made by a subject were true and half were false. An interviewer sat across a table from a subject and on each of 40 trials he held up a 3" x 5" card with either a word or a trigram printed on it. He then asked "Did you cross out this item?". A number between 1 and 200 was printed in the top right-hand corner of each card and, as in

Experiment 2, subjects were given a simple code for determining their verbal response (either "Yes I did", or "No I did not"). No rehearsal procedure was employed in this experiment. Instead, subjects were instructed to make their response as soon as possible after the interviewer finished asking a question. As before, subjects were instructed to look right at the interviewer when making their response.

In order to obtain the recall data, the following cover story was given to subjects:

Again, the interviewer will be trying to make a judgment about the truthfulness of your answer, and his judgments should be influenced by the way you answered his particular questions earlier on. One final thing. I want the interviewer to think that you are making a judgment as to whether or not you think he believed the statement you just made. I want him to think that you are trying to guess how he judged your answer. So, I want you to make a decision after each answer that you give. After you answer the interviewer, he will put the card with the word or trigram on it on the table. At that point, I want you to try to recall whether you did or did not cross out that item earlier (A subject was then given a small plastic box with two slots in the lid, marked DID and DID NOT. He was instructed to file each stimulus card according to his memory in one of the two slots. In the test room the box was placed in an open drawer below the top of the table so that the interviewer could not observe how a subject filed the stimulus cards.). Although you may find this difficult, just try to do the best you can and file each card in one of the two slots. This will give the interviewer the impression that you are making a judgment about what he is thinking, and also the memory data that it produces can be used.

Twenty of the 40 stimulus items, 10 words and 10 trigrams, were employed in each of the four true statement conditions: truth interviewer A, lie interviewer B, truth interviewer B, and lie interviewer A. The remaining 20 stimulus items, 10 words and 10 trigrams, were employed in each of the

four false statement conditions: truth interviewer A, lie interviewer B, truth interviewer B, and lie interviewer A. Half of the stimulus items had actually been crossed out by a subject; half had not been crossed out. There were no control conditions. This procedure yields a complete $2 \times 2 \times 2 \times 2$ between-within-subjects design with ten observations per cell on the dependent measure, recall accuracy.

Results and Discussion

In order to make the presentation manageable, the data is broken down according to the interviewer's identity. Table 10 presents the mean proportions of hits and false alarms, as well as the corresponding d' values for each of the eight experimental conditions in which interviewer A served as a credibility cue. Table 11 presents these same statistics for the corresponding experimental conditions in which interviewer B served as a credibility cue. The d' values are also plotted for each experimental condition in Figures 6 and 7. An overall analysis of variance performed on the d' data is summarized in Table 12. It indicates a main effect for the stimulus variable which approaches conventional significance levels ($F=3.68$, $df\ 1/36$, $p<.07$). This suggests that the manipulation of internal cue distinctiveness was successful. Recall performance was poorer for trigram operations ($\bar{X}=0.51$) than for word operations ($\bar{X}=0.90$). The analysis also indicated a significant main effect for the interviewer identity variable ($F=5.11$, $df\ 1/36$, $p<.05$), and the absence of main effects for the interviewer cue and statement variables. This suggests that the two interviewers

Table 10. Mean proportions of hits, false alarms, and corresponding d' values for recall of word and trigram operations in each statement and interviewer cue condition for interviewer A.

Statement Condition		Stimulus Condition			
		WORDS		TRIGRAMS	
		Interviewer Cue		Interviewer Cue	
		TRUTH	LIE	TRUTH	LIE
	Hits	.80	.80	.59	.39
TRUE	False Alarms	.65	.69	.18	.42
	d'	0.62	0.42	1.52	0.01
	Hits	.72	.72	.34	.46
FALSE	False Alarms	.64	.56	.34	.28
	d'	0.11	0.54	0.01	0.68

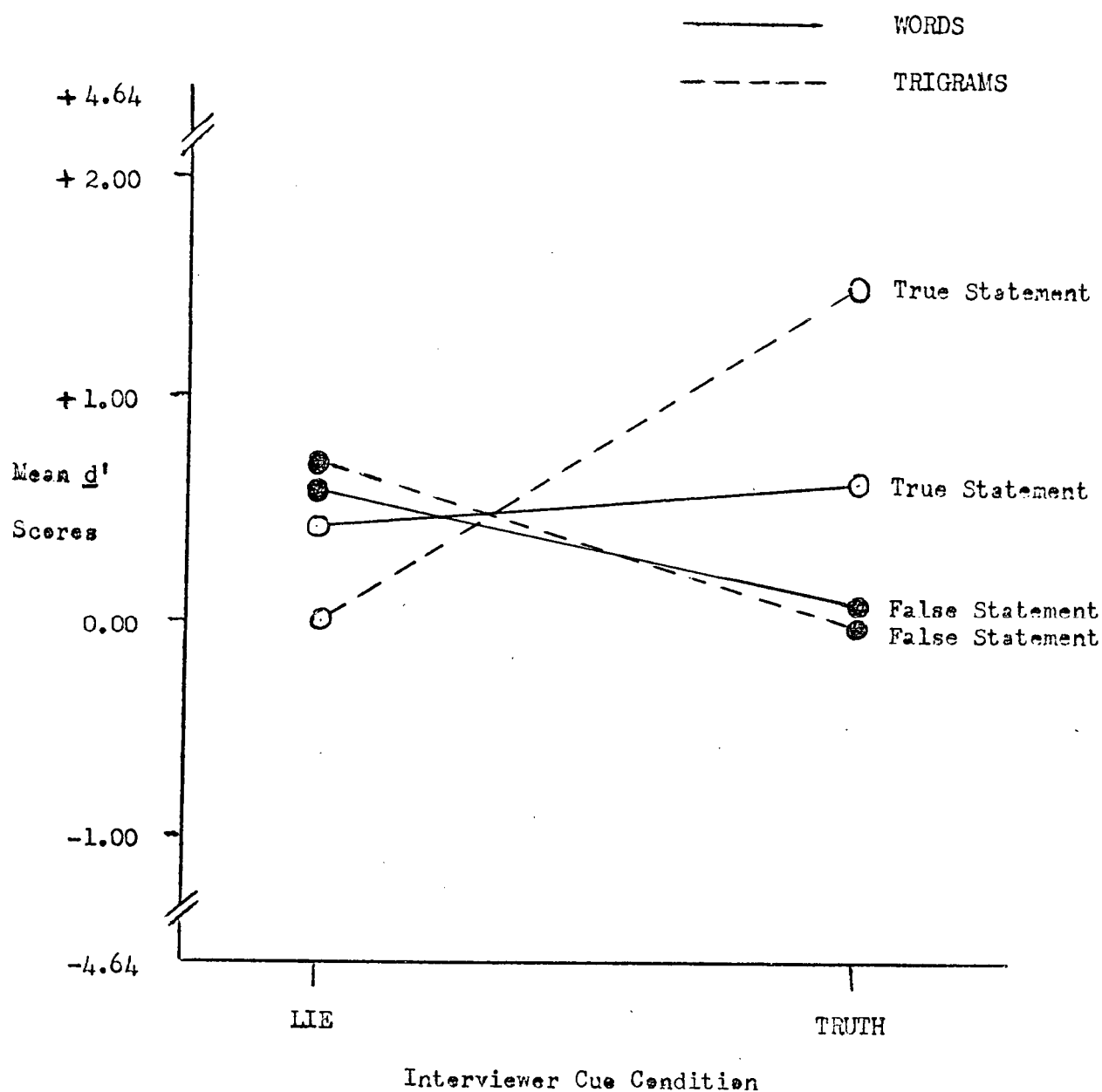


Figure 6. Recall accuracy for word and trigram operations in each statement and interviewer cue condition for interviewer A.

Table 11. Mean proportions of hits, false alarms, and corresponding d' values for recall of word and trigram operations in each statement and interviewer cue condition for interviewer B.

Statement Condition		Stimulus Condition			
		WORDS		TRIGRAMS	
		Interviewer Cue		Interviewer Cue	
		TRUTH	LIE	TRUTH	LIE
	Hits	.80	.86	.41	.58
TRUE	False Alarms	.52	.43	.51	.40
	d'	1.24	1.66	-0.28	0.57
	Hits	.68	.86	.49	.45
FALSE	False Alarms	.35	.48	.24	.28
	d'	1.21	1.43	1.14	0.46

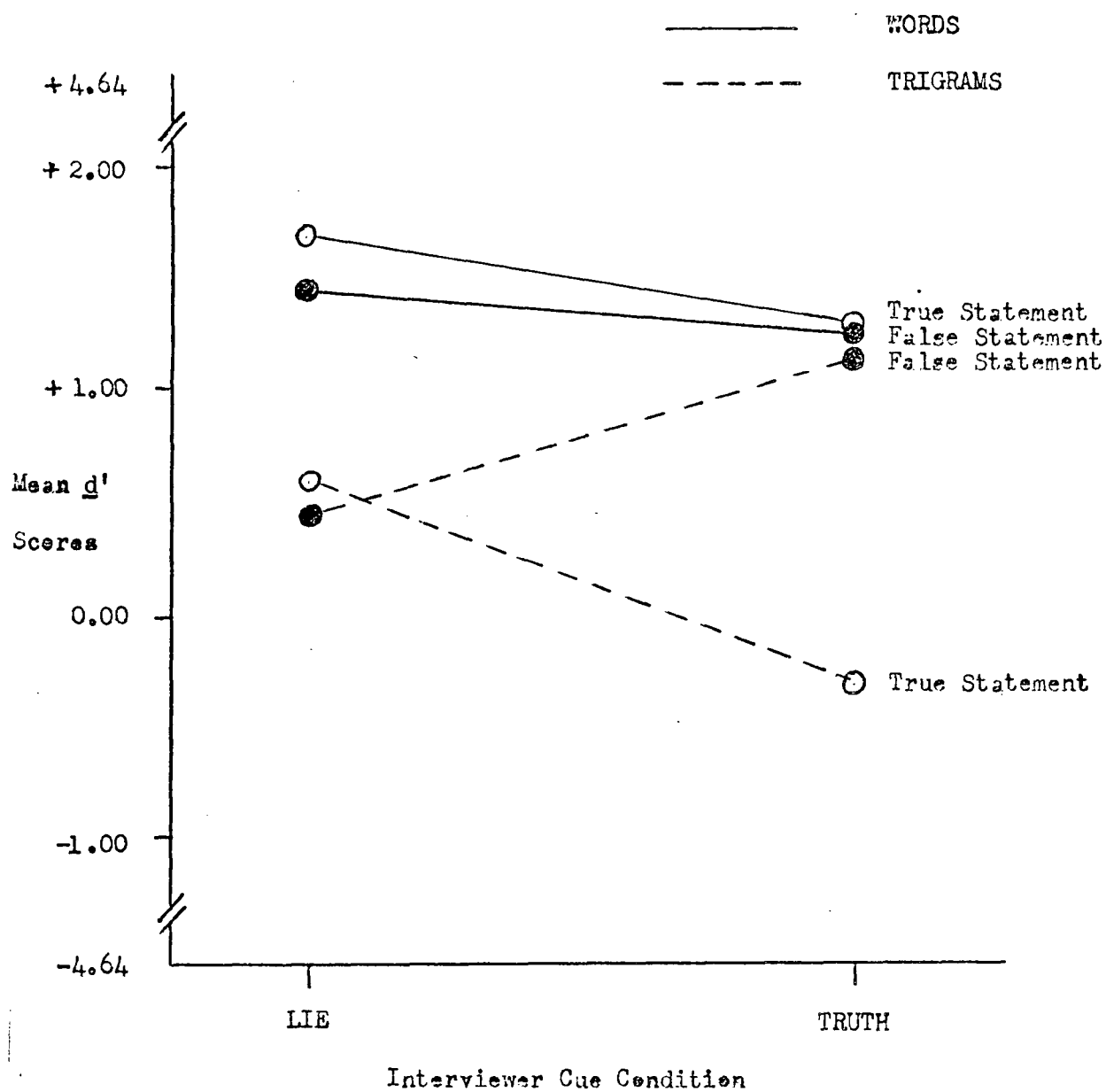


Figure 7. Recall accuracy for word and trigram operations in each statement and interviewer cue condition for interviewer B.

Table 12. Analysis of variance of recall scores (d').

Source	df	MS	F
Total	159		
<u>Between</u>	<u>29</u>		
Interviewer Cue (A)	1	0.04	0.03
Interviewer Identity (B)	1	7.71	5.11**
A x B	1	1.25	0.83
Subjects within groups	36	1.51	
<u>Within</u>	<u>120</u>		
Statement (C)	1	0.02	0.01
Stimulus (D)	1	6.11	3.68*
C x D	1	0.77	0.46
C x A	1	0.68	0.43
C x B	1	3.20	2.03
D x A	1	1.46	0.88
D x B	1	10.94	6.59***
C x A x B	1	14.43	9.19****
D x A x B	1	0.21	0.13
C x D x A	1	0.06	0.04
C x D x B	1	2.26	1.35
C x D x A x B	1	3.98	2.37
C x Subjects within groups	36	1.57	
D x Subjects within groups	36	1.66	

Table 12. (cont.)

Source	df	MS	F
C x D x Subjects within groups	36	1.68	

* $p < .07$

** $p < .05$

*** $p < .025$

**** $p < .005$

had different effects upon recall accuracy. This effect should not be confused with possible differential credibility cueing effects by the two interviewers. It simply indicates that recall performance was poorer in the presence of interviewer A ($\bar{X} = 0.49$) than in the presence of interviewer B ($\bar{X} = 0.93$).

Two significant interactions emerged from the analysis. These interactions were between the stimulus and interviewer identity variables ($F = 6.59$, df 1/36, $p < .025$), and between the interviewer identity, statement, and interviewer cue variables ($F = 9.19$, df 1/36, $p < .005$). The significant two-way interaction, shown in Figure 8, suggests that the main effect for stimuli was entirely due to interviewer B. Because the required comparisons were incidental or post hoc, a Newman-Keuls test was employed to confirm this observation. Recall accuracy was significantly higher for words ($\bar{X} = 1.39$) than for trigrams ($\bar{X} = 0.47$) in the presence of interviewer B ($p < .01$). There was no difference in recall accuracy for words ($\bar{X} = 0.42$) and trigrams ($\bar{X} = 0.56$) in the presence of interviewer A. The significant three-way interaction, shown in Figure 9, suggests that only one of the two interviewers obtained the predicted credibility cueing effect. Because the comparisons required here were planned in advance (recall that interviewer identity was manipulated as an independent variable in order to test for differential credibility cueing effects), two orthogonal linear contrasts were employed to confirm this observation. The expected interaction between interviewer cues and statement cues was significant for interviewer A ($F = 6.29$, df 1/36, $p < .025$) but not

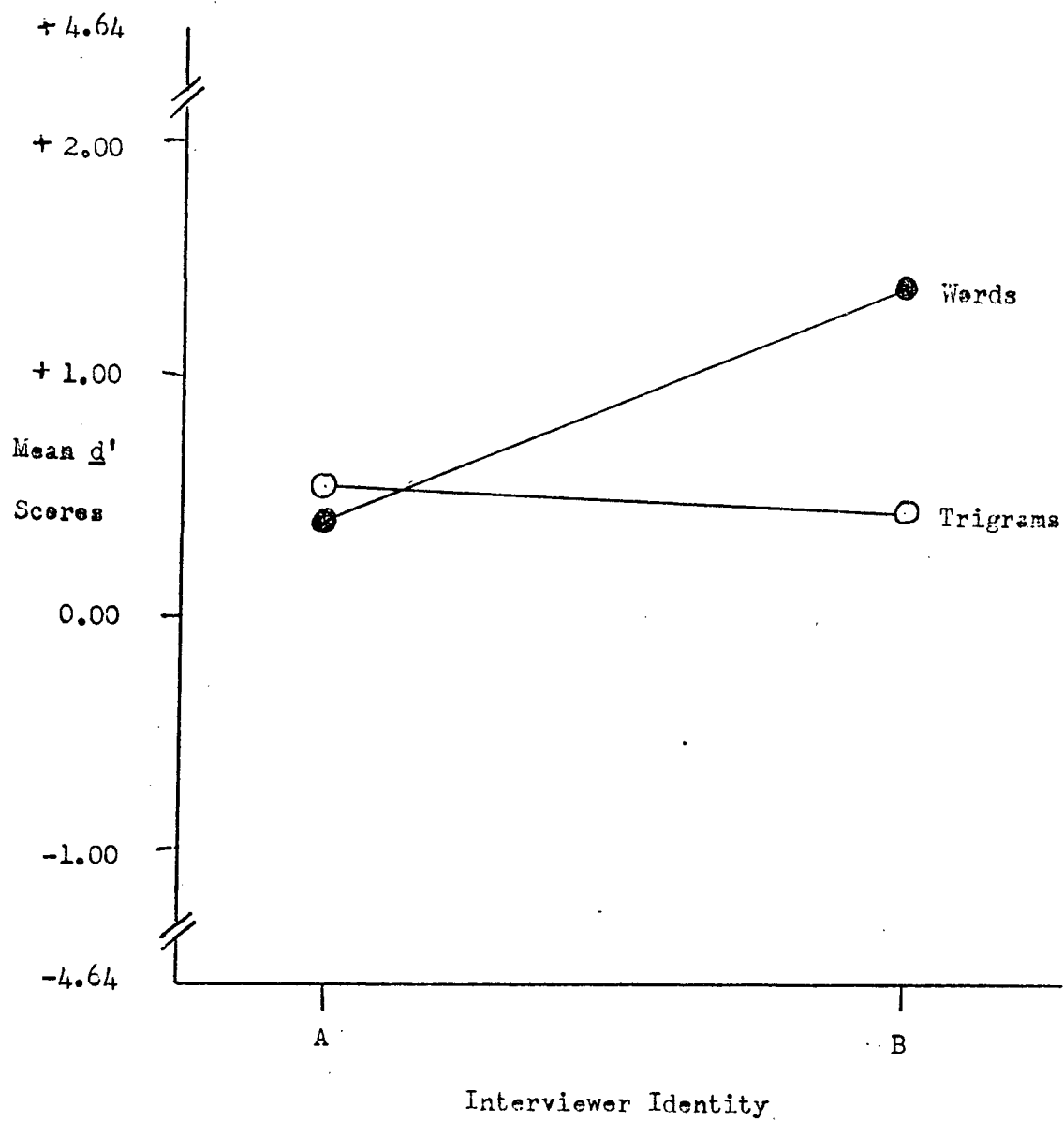


Figure 8. Recall accuracy for words and trigrams in each interviewer identity condition.

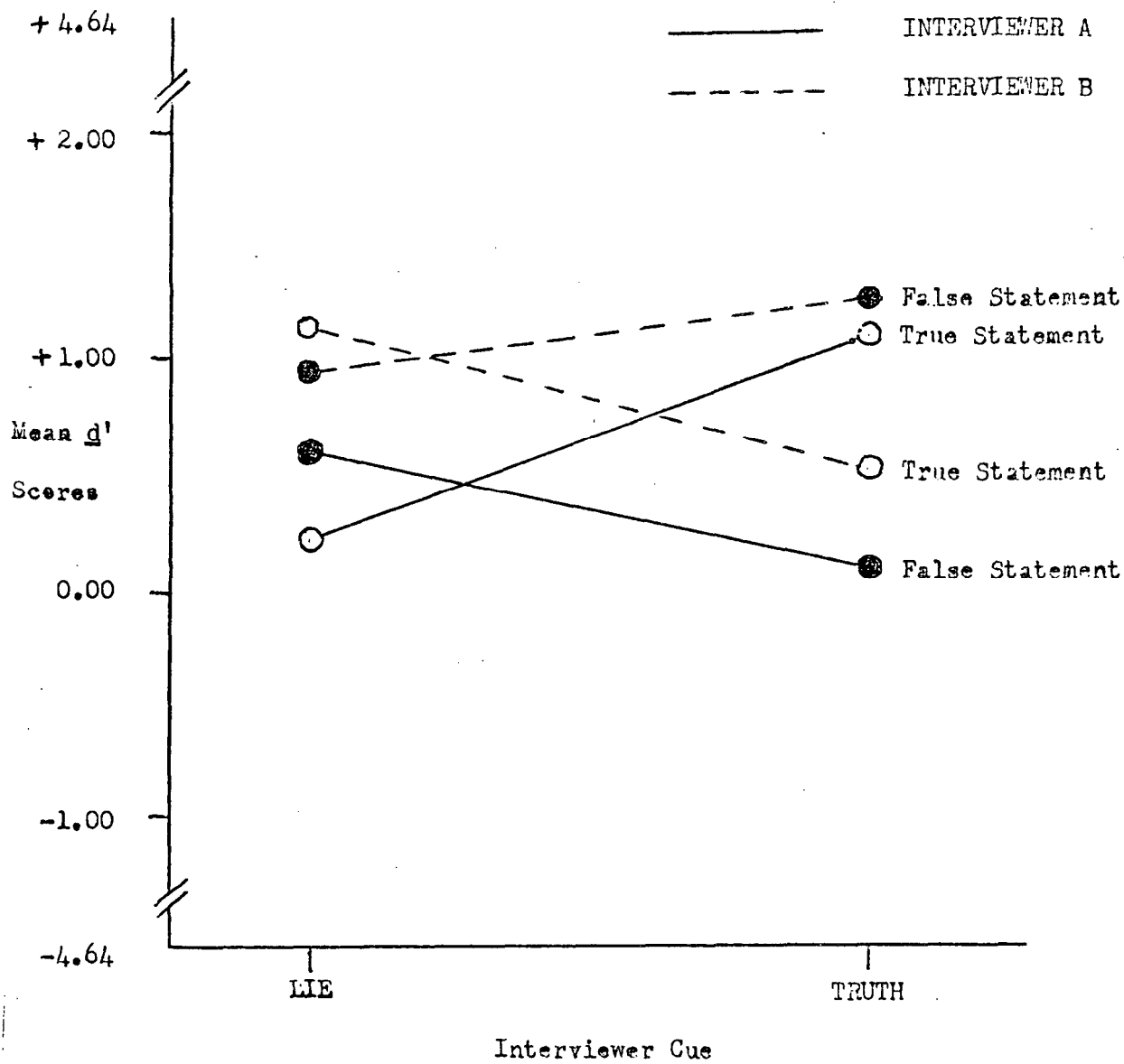


Figure 9. Recall accuracy following true and false statements in each interviewer cue and interviewer identity condition.

for interviewer B ($F=2.83$, df 1/36). The significant interactive effect of interviewer cues and statement cues for interviewer A is consistent with the general prediction of poorer recall performance in those conditions where these two cues provide misleading information. The mean d' value for the combined false statement-truth interviewer and true statement-lie interviewer conditions is 0.14 as compared with a mean d' value of 0.84 in the combined true statement-truth interviewer and false statement-lie interviewer conditions.

The data for interviewer A were analyzed further in order to examine the hypothesized components of the external cue effect. The results provide support for both of the hypothesized components. False statements made in the presence of the truth interviewer cue resulted in poorer recall performance than did false statements made in the presence of the lie interviewer cue ($\bar{X}s=0.07$ and 0.61 respectively; $t=2.03$, df 18, $p<.06$).¹¹ Similarly, true statements made in the presence of the lie interviewer cue resulted in significantly poorer recall performance than did true statements made in the presence of the truth interviewer cue ($\bar{X}s=0.21$ and 1.07 respectively; $t=2.33$, df 18, $p<.05$). The self-perception interpretation of these results has been outlined earlier (see p. 79). A glance at Figure 6 suggests that these effects are more prominent in the recall of trigram operations than word operations, as would be expected on the basis of the internal cue distinctiveness hypothesis. Given the overall pattern of results,

¹¹ All t -tests are 2-tailed.

a significant 4-way interaction would provide statistical support for this hypothesis. Since this interaction failed to reach conventional significance levels, such support is lacking. Because there is no statistical evidence that the internal cue distinctiveness manipulation was successful for recall in the presence of interviewer A (no differences in recall accuracy for words and trigrams), however, this result is not surprising.

While recall performance in the presence of interviewer A lends itself to a self-perception interpretation in a fairly straight-forward manner, recall in the presence of B is puzzling. Although Figure 9 indicates that this latter pattern of data is the reverse of what would be expected on the basis of the self-perception theory, statistical analyses provide some measure of comfort for the theory. As was previously noted, the linear contrast which examined this "reverse" interaction was insignificant. As this result would suggest, subsequent 2-tailed comparisons corresponding to those made for interviewer A also failed to approach conventional significance levels. There were no differences in recall performance between the false statement-truth interviewer cue and false statement-lie interviewer cue conditions ($\bar{X}s = 1.18$ and 0.94 respectively; $t = 0.67$) and between the true statement-lie interviewer cue and true statement-truth interviewer cue conditions ($\bar{X}s = 1.12$ and 0.48 respectively; $t = 1.33$).

When the significant stimulus variable is taken into consideration in examining the data for interviewer B (shown in Figure 7), three final comparisons are suggested. These involve recall performance for

trigram operations. When only the trigram data were subjected to the same 2-tailed comparisons described above for word and trigram data combined, the results were unchanged. Again, there were no significant differences in recall performance between the false statement-truth interviewer cue and false statement-lie interviewer cue conditions ($\bar{X}_s=1.14$ and 0.46 respectively; $t=1.33$) and between the true statement-lie interviewer cue and true statement-truth interviewer cue conditions ($\bar{X}_s=0.57$ and -0.28 respectively; $t=1.49$). However, recall of trigram operations was significantly poorer in the true statement-truth interviewer cue condition than in the false statement-truth interviewer cue condition ($\bar{X}_s=-0.28$ and 1.14 respectively; $t=3.23$, $df\ 9$, $p<.02$). This finding is contrary to the self-perception predictions and the reverse of the statement cueing effect obtained by the truth interviewer in Experiment 2. The interpretation of these results was that cues previously associated with truth telling induced subjects to believe verbal statements made in their presence. Hence, a related interpretation of the present trigram results for interviewer B would have to hold that cues previously associated with truth telling induced subjects to disbelieve verbal statements made in their presence. Although such an interpretation makes no sense in terms of the self-perception theory, the two data points which render it possible are not easily explained away. For example, it would be tenuous to argue that the statistical significance of this comparison was due to chance while taking the position credibility cueing was responsible for the statistical significance of the corresponding comparison in the previous experiment

(For a discussion of difficulties in interpreting the results of multiple t-tests, see Hays, 1963, pp. 471-472.).

A simple interpretation of the present results is that they provide support for the notion that some stimulus persons are more likely to produce credibility cuing effects than others. In this experiment the effect was produced by one interviewer but not by the other. How should this be explained? A clue is suggested by the fact that overall recall performance was poorer in the presence of the interviewer who produced the predicted cuing effect. This observation is consistent with the internal cue distinctiveness hypothesis derived from the self-perception theory. By employing stimuli designed to produce internal cues differing in distinctiveness it is presumed that this hypothesis was operationalized at the storage end of a memory process (i.e., trigram operations should produce weaker internal cues than word operations, presumably in the form of memory traces).¹² Of course, this is not the only way in which internal cues might be rendered more or less distinct. Operations performed at the retrieval end of a memory process could have a similar effect. For example, the introduction of a time limit on recall might be expected to render internal cues less distinct, or less accessible, than they would be in the

¹²At the level of behavior, distinctiveness is clearly a statistical concept (more word operations than trigram operations recalled correctly). Although the terminology of the distinctiveness hypothesis implies that distinctiveness is a statistical concept at the mediational level as well (more word operations than trigram operations stored in memory), the author is committed to this notion for the sake of terminological consistency only. It does not reflect a theoretical commitment.

absence of this constraint. Such a procedure would have a debilitating effect upon recall in general, and, according to the internal cue distinctiveness hypothesis, it should bring recall more under the control of available external cues.

There is some evidence that a time-limited, retrieval-based operationalization of internal cue distinctiveness was inadvertently produced in the present experiment. This was likely due to the casual way in which the interviewer identity variable was manipulated. Because there were no especially obvious interpersonal characteristics and behaviors which might produce the credibility cueing effects, none were selected for manipulation in this experiment. That is, no attempt was made to have one interviewer behave in a particular manner that might be considered effective in producing cueing effects, while having the other interviewer behave differently. Instead, interviewers were allowed to interact with subjects relatively freely, particularly during the test phase of the experiment. Although both interviewers were given the same standard routine to follow, each was allowed to proceed at his own pace and level of formality when presenting the test stimuli to a subject. One observable consequence of this arrangement was that interviewer A consistently presented the test stimuli at a faster pace than interviewer B. Although both interviewers began the test phase of the experiment at approximately the same time, the interviewer who produced credibility cueing effects was always the first to finish with his subject. If it can be assumed that this rapid pace was imposed upon the time of recall testing as well as

stimulus presentation, then a possible effect may have been to weaken the presumed storage-based manipulation of internal cue distinctiveness. Since recall of trigram operations is already marginal, the imposition of a rapid pace would have its greatest effect upon recall of word operations. A time constraint should have a debilitating effect upon the retrieval of otherwise accessible information. This would bring recall of word operations closer to the marginal level of recall displayed for trigram operations than would occur with a more leisurely pace. The same low level of recall accuracy for both word and trigram operations in the presence of interviewer A, and the interactive effect of the statement and interviewer cue variables for both of these stimulus operations combined in the presence of A, are consistent with this speculation.

What appears to have occurred in the present experiment can be summarized in terms of the presumed storage- and retrieval-based manipulations of internal cue distinctiveness. The presumed storage-based manipulation of internal cue distinctiveness was successful for subjects who attempted to recall in the presence of interviewer B (recall accuracy poorer for trigram operations than word operations). Due to the relatively rapid pace imposed by interviewer A, this manipulation had no effect for subjects who attempted to recall in his presence (no difference in recall accuracy for word and trigram operations). As a result, instead of the intended experimental manipulation of internal cue distinctiveness based upon storage, there was an inadvertent interviewer manipulation of internal cue distinctiveness

based upon retrieval. Rather than establishing internal cues at two levels of distinctiveness for subjects within the interviewer identity variable (storage of trigram operations poorer than storage of word operations), internal cues were established at two levels of distinctiveness for subjects between the interviewer identity variable (retrieval of word and trigram operations combined in the presence of interviewer A poorer than the retrieval of these operations in the presence of interviewer B).

While this explanation would account for the greater likelihood of credibility cueing effects in the presence of interviewer A, it does not explain the absence of these effects in B's presence. Since the presumed storage-based manipulation of internal cue distinctiveness was successful for subjects attempting to recall in the presence of B, and since recall of trigram operations was at the same low level of accuracy in the presence of both interviewers, recall of trigram operations in B's presence would be expected to show these effects. An impressionistic speculation is offered as a tentative explanation for their absence. It is my impression that the two individuals employed in this experiment differed strikingly in their tolerance for the "interviewer" role. As it was described to these two individuals, the role of an interviewer was formal, perfunctory, and rigid. Interviewers were instructed to ask a predetermined set of questions according to a dry and repetitive format, to stare directly at subjects in order to give the appearance of evaluating their answers, and to attempt to suppress their own reactions to subjects' answers in order

to minimize the effects of extraneous, non-verbal cues. This procedure was repeated in 20 sessions. Interviewer A appeared to be quite comfortable with this role and was able to play it well. My own impression of his performance is that he was "cool", "efficient", "serious", and "mechanical". He appeared to transmit very few interpersonal cues other than those required by his particular role as the party controlling an interaction. Interviewer B, on the other hand, appeared to be uncomfortable with this role and as a result he behaved differently. My impression of his performance is that he was "warm", "sensitive", "cavalier", and "human". In contrast to the somewhat "bureaucratic" performance of A, B carried off his role with considerable animation and gave the appearance of being sensitive to subjects as individuals. He tended to transmit interpersonal cues which were inconsistent with his role as the party controlling an interaction. Allowing subjects to proceed pretty much at their own pace in the recall phase of the experiment is a concrete example of this tendency. His insistence that the "interviewers" be allowed to talk with subjects in order to divest themselves of their roles is another. B specifically requested that the interviewers be allowed to assure subjects that they were not really the "probing, mechanical zombies" that their roles made them out to be. In short, subjects would likely form relatively differentiated impressions of the kind of person B is much more readily than they would for A, an assumption derived from the correspondent inference theory of Jones and Davis (1965). That theory holds that out-of-role behavior provides an informational basis for launching

inferences about an actor's "personality characteristics", whereas in-role behavior provides little or no information for such inferences.¹³

Given that B's behavior resulted in the transmission of considerable information about his "personality", a possible effect of this personality cueing would be to obscure the credibility cues which the experiment was designed to establish. Given that A's behavior resulted in the transmission of very little of this kind of information, credibility cues would not be obscured in his presence by personality cues. Taken together, these differential personality and credibility cueing effects could produce the general pattern of results observed in the present experiment. Since no impression formation data were collected from subjects, however, this explanation is merely speculative.

Credibility cueing in a non-laboratory context: Bem (1970) has speculated that the credibility cueing procedure, with particular reference to "truth" cueing, could have important implications in the

¹³The notion that considerable information is extracted from out-of-role behavior is supported in an experiment conducted by Jones, Davis, and Gergen (1961). Subjects were exposed to tape-recorded interviews in which individuals behaved in either an "inner-directed" or "outer-directed" manner in applying for a job which required either an "inner-directed" person (astronaut) or an "outer-directed" person (submariner). When asked what they thought the applicant was really like as a person, subjects exposed to the out-of-role recordings gave more extreme ratings to the applicant and expressed greater confidence in their judgments. The outer-directed astronaut (out-of-role) was seen as very affiliative and very conforming whereas the inner-directed astronaut (in-role) was seen as moderately independent and moderately affiliative. Similarly, the inner-directed submariner (out-of-role) was seen as very independent and very non-affiliative whereas the outer-directed submariner (in-role) was seen as moderately affiliative and moderately conforming.

area of police interrogations. Specifically, he has presented his findings to the U.S. Senate Subcommittee on Constitutional Amendments and argued that interrogation procedures might have dangerous psychological effects on the beliefs and memories of an innocent person who is suspected of committing a crime. According to this argument, elements of the credibility cueing procedure may be inherently present, or at least available, during an actual police interrogation. For example, the physical surroundings of any police interrogation may act as truth cues for an innocent suspect because he simply could not conceive of himself lying in such circumstances. As a result, he is in a situation in which he is likely to be influenced by his own statements. Any inadvertent errors or inaccuracies he makes in an otherwise truthful testimony are likely to be accepted thereafter by him as true. In addition to this inherent predisposition toward self-credibility, a skilled police interrogator can elicit inaccuracies and distortions by asking leading questions or by carefully programming the sequence of questions asked. These inaccuracies could then come to be believed by the suspect in much the same way that experimental subjects perceived their false statements to be truthful. Finally, Bem suggests that the recently imposed U.S. Supreme Court guidelines for interrogations may function as truth cues facilitating belief in any inadvertent errors a suspect might make in his statements. That is, if the police can elicit false incriminating statements from a suspect after they have informed him of his right to remain silent and his right to counsel, these statements are likely to be believed

by the suspect thereafter since they cannot be attributed by him to coercion. Such statements should be perceived by the suspect as given of his own free will, in a situation where he knows such statements can be subsequently used against him in a court of law, and therefore signal him that he is telling the truth. These speculations lead Bem to an ironic conclusion: it seems that the less an interrogator uses coercive tactics and the more he takes advantage of truth cue contingencies inherent in the interrogation situation, the more susceptible a suspect becomes to thought control through self-perception (Note that witnesses would also be susceptible to the influence of some of these same cues.).

The present results could be applied to this argument in the interests of a more complete speculation. While Bem has drawn attention to aspects of the interrogation situation which could cause a suspect to believe his own verbal inaccuracies, he has not addressed in any detail the important question of how errors in testimony might occur in the first place. It is one thing to get experimental subjects to make a preprogrammed set of false statements concerning a trivial task; it is quite another thing to induce errors in a suspect's testimony concerning a crime. An examination of official police interrogation manuals is instructive on this point. In each of three manuals considered (Inbau and Reid, 1953; Aubry and Caputo, 1965; O'Hara, 1970), a peculiar psychology of the interrogation process is described. The central assumption of this psychology is expressed by the adage, "Truth will out". According to this view, people have a basic need to tell the truth, even when it is not in their best interests to do so, and a

skillful interrogator is one who can create conditions which facilitate the telling of such truths by persons who have committed crimes. A corollary to this "rule" is the assumption that an innocent suspect cannot be led to falsely incriminate himself. This psychology, the style of its supportive logic, and the priority of securing confessions are clearly expressed in the following passage from an interrogation manual:

The basic purpose of an interrogation is the securing of an admission of guilt from a person who has committed a crime We might mention here the all-important fact that the correctly prepared and correctly structured interrogation is not likely to produce an admission of guilt and a confession from an individual who is not guilty of the commission of the crime for which he is being interrogated. The reason for this is basic and is firmly anchored in the Constitution of the United States; and upon the fact that every citizen of the United States has Constitutional rights, God given and inalienable, which flow from and are guaranteed by the Constitution, and which may not be taken away from the individual.

(Aubry and Caputo, 1965, p. 25).

Presumably, it is this peculiar psychology which legitimizes a standard set of interrogation tactics advocated in each of the manuals. A partial list includes "The Mutt and Jeff Technique", "Reverse Line-Up", "Bluff on Split Pair", "Teed Down-Build Up Technique", and "What Will the Wife and Kids Think?". In the "Mutt and Jeff Technique", for example, one interrogator behaves in a friendly fashion toward the suspect while the other is angry and threatening. When Jeff is out of the interrogation room, Mutt confides that he doesn't think he can hold off his violence-prone associate much longer, and that it would be "best" for the suspect to confess before Jeff gets out of control.

In the "Reverse Line-Up", the suspect is identified by a confederate posing as a witness. In addition to these orchestrated plays, interrogators are instructed to sit close to the suspect so as to make him feel uncomfortable, to create an illusion of psychological isolation, to display confidence in the suspect's guilt, not to allow him a chance to explain or issue repeated denials, to semantically play down the seriousness of the offense, to persist and give the impression of super-human endurance, to note symptoms of lying and point these out to the suspect (sweating, inability to look the interrogator in the eye, pulsating carotid artery, long pauses), to advise the suspect of his right to remain silent and then point out to him the implications of his doing so (guilty), etc.

Almost all of these tactics are based upon deception and their stated goal is to produce tension and confusion on the part of the suspect. It is assumed that when the suspect is in this disadvantaged state the "truth will (pop) out". Consider the following observations to this effect:

(The interrogator must be able to) create the impression in the mind of the suspect that no undue advantage is going to be taken of him that he is not going to be tricked or outmaneuvered into making the incriminating admission the interrogator must possess a bag of tricks similar to the top hat of a magician, and like the magician who pulls rabbits and other items out of the top hat, the interrogator must be able to pull various approaches out of his bag of tricks, and do so with the ease and competence that only comes with long hours of practice, experimentation, and application The use of subterfuge makes a very effective approach, and occasionally may be used with telling force and effect. In a certain sense, subterfuge may be thought of as trickery, although it is not defined as trickery. (Aubry and Caputo, 1965, pp. 76-78).

The suspect is in a state of emotional confusion. He is unable to think logically and clearly, since his sense of values has been disturbed and his imagination is distorting perspective. It is possible for the investigator to obtain admissions or even a confession by further misrepresenting the picture.

(O'Hara, 1970, p. 821).

When the interrogator initiates the interrogation with gentleness and kindness, instead of the expected roughness, the net result on the part of the subject is going to be extreme confusion. The confusion will work to the advantage of the interrogator, and the subject may well confess before he even realizes what he is saying.

(Aubry and Caputo, 1965, p. 81).

.... subjects accused of a crime; apprehended and charged with it; and particularly while actively being interrogated for it; are not apt to be using the intellectual powers of their mind to any great extent. If any of their mental faculties are functioning at all, they will most likely be in the psychological and emotional areas. An individual in a nervous, upset, and highly emotional state is very amenable to suggestions, and will tend to comply with and carry out suggestions that are made to him; much in the same manner that conditioned reflexes are carried out.

(Aubry and Caputo, 1965, pp. 114-115).

.... the interrogator must give no indication that he is being influenced by what the subject may state in behalf of his innocence; and this should be so even when the interrogator actually realizes the reasonable implication of possible innocence in some fact or evidence presented by the subject. In other words, the subject should be required to extend himself to the limit in order to avoid detection or confession, for during the course of his efforts toward that end he is more vulnerable to the tactics and techniques designed to produce the desired information.

(Inbau and Reid, 1953, p. 153).

The "tone" of the interrogation is set by the interrogator. He may give the impression that he is completely convinced of the subject's guilt, depending upon the facts of the case, in which instance the subject will have to extend himself to convince the interrogator that he is wrong. And it is

just at the moments that the subject is extending himself the most, that he is most likely to make serious errors. (Aubry and Caputo, 1965, p. 153).

While the stated purpose of these tactics and procedures is the securing of the right information from suspects who are factually guilty, an inadvertent effect of the tension and confusion they are designed to produce might well be the securing of the wrong information from suspects who are factually innocent. The well documented interrogation of George Whitmore by the New York police is a case in point (Shapiro, 1970). Subjected to many of the confusing and suggestive interrogation tactics just described, Whitmore confesses to three murders which he did not commit.¹⁴ In two important respects Whitmore appears to be a classic case for the application of self-perception theory to police interrogations. He was possessed of a chronically poor memory (possibly predisposed toward storage in the form of weak internal cues), a disposition which was exaggerated by confusing and suggestive interrogation tactics (further weakening of internal cues in retrieval). One effect of the combination of weak internal memory cues and confusing interrogation tactics was that the police were able to elicit incriminating errors, including false confessions, in their interrogation of Whitmore. A second effect of the combination of weak internal memory cues and blatantly suggestive interrogation tactics was that Whitmore repeated

¹⁴ After spending almost two years in various penal and mental institutions, Whitmore was cleared of any responsibility for the three murders to which he confessed. His case has been cited as a decisive factor in the U.S. Supreme Court's 196 decision to limit, at least technically, the powers of the police in interrogating suspects.

in court many of the same incriminating errors he had earlier made to the police.¹⁵ Although limited in scope, the present experiment can be viewed as an analog to this particular police interrogation situation. The presumed weakening of retrieval-based internal cues by interrogator A's fast pace is analagous in principle to the weakening of internal cues by confusing police interrogation tactics. The preprogrammed set of false statements concerning the crossing out task is analagous to the erroneous statements suggested by the police interrogators. A subject's belief in the accuracy of some of his false statements in the recall test is analagous to Whitmore's belief in the accuracy of the erroneous parts of his testimony offered in court. Although Whitmore subsequently revoked his confessions, he persisted in making serious incriminating statements in his testimony in court, statements which could not possibly be true. It seems quite likely that Whitmore came to believe these statements in much the same way experimental subjects believed their false statements ----- by processing information from his own verbal behavior and the situation in which it occurred.

Finally, the present results suggest that a particular kind of interviewer/interrogator is more likely than others to obtain credibility cueing effects. This individual is someone who transmits little information about his "personality characteristics", perhaps in the manner of the inscrutable Sergeant Friday of the television series, Dragnet ("Just the facts, that's all we want, just the facts.").

¹⁵As an example of suggestive tactics, by giving "clues" an interrogator was able to get Whitmore to draw an accurate and detailed floor plan of the apartment of an alleged victim, even though he had never seen it.

CHAPTER FOUR: SUMMARY AND DISCUSSION

This dissertation took up a new "mini-paradigm" in Social Psychology, more or less on the bases of faith and aesthetics, and attempted to develop some hard-headed support for it. If the present research has articulated the self-perception theory further than had previously been accomplished by Bem and his detractors, this metatheoretical task will have met with some success. By way of examining this contingency, the final chapter summarizes the experimental findings and considers their implications for the theory.

In testing a theory, the results of any intervention which increases the subjective likelihood that an experimental effect is mediated by a process postulated by the theory take precedence over other results of that experiment. For this reason the elimination of a differential guessing interpretation of cueing data by use of the d' statistic must be viewed as the most important accomplishment of the first experiment. Ruling out the possibility of this artifact increases confidence in a credibility cueing interpretation of the recall data. In effect, the internal validity of results produced by Bem's procedure has been established more firmly than it had been prior to this experiment.

A second accomplishment of the first experiment is the provision of support for a fundamental hypothesis derived from the self-perception theory. It was observed that external cues provided by subjects' verbal behavior and the discriminative light stimuli had their effect in the weak internal cue condition and not in the relatively stronger internal cue condition. Because effects of different magnitude were predicted in

the two internal cue conditions (large effect for the weak condition and smaller effect for the stronger condition), support for the internal cue distinctiveness hypothesis was viewed as "qualified". It should be noted, however, that this qualification is based upon empirical precedent rather than theory. Whereas Bem has previously obtained a credibility cueing effect for recall using word stimuli, no such effect was observed in this experiment. Yet the theory of self-perception simply assumes that credibility cueing effects are more likely under conditions of low internal cue strength than under conditions of high internal cue strength. The observation of an effect in the recall of trigram operations but not in the recall of word operations is consistent with this assumption. Nevertheless, given Bem's findings, a question is raised by the present failure to obtain a credibility cueing effect in the recall of word operations. This question concerns the robustness of the credibility cueing procedure. Of three experiments using this procedure, only one has obtained the interactive effect of behavioral and situational cues for recall of word operations (Bem, 1966). Maslach (1971) observed a main effect for situational cues (better recall in the lie light conditions) and no cueing effect was observed in the present experiment. While Maslach's claim to have found an "alternative explanation" of Bem's findings is derived from an eccentric interpretation of the term "artifact", and should therefore be dismissed as foolishness, the actual results of her experiment are important as they relate to the question at hand.¹⁶ Along with the word results of

¹⁶On the basis of her results Maslach claims that the "truth about

the present experiment, these findings suggest that the credibility cueing procedure is rather fragile so that unless certain conditions are met the results predicted by the self-perception theory will not be forthcoming. At present these conditions have not been specified in any other than gross procedural terms. In this connection, the cueing effect was not observed in a pilot study in which a slide projector was used to present the recall stimuli. It may be that any manipulation which diverts the subject's attention from the behavioral and situational stimuli, such as the projection of recall stimuli in a visual medium, washes out the effect. A manipulation which directs subjects' attention to one of the discriminative situational stimuli at the expense of the other, as seems to have occurred in Maslach's

(belief in) false confessions" is that they are the product of subjects' decreased "vigilance" in the presence of the truth light. While this could explain the pattern of data she obtained (a main effect for light stimuli), it cannot explain Bem's results (an interactive effect of statement and light stimuli). In fact, Bem's observation of relatively accurate recall in his true statement-truth light condition and relatively inaccurate recall in his true statement-lie light condition contradicts the proposed "vigilance" explanation. The observation of this same interaction for recall of trigram operations in the present experiment reduces the tenability of this explanation still further. This experiment included a procedural control designed to rule out a possible "vigilance" effect. The discrepancy between the findings of Bem and Maslach suggests the possible operation of some artifact in the latter experiment rather than the operation of a "vigilance" artifact in the former. Maslach's findings may well have been produced by her use of a polygraph. By emphasizing lie detection to her subjects in this way, she may have differentially sensitized them to lying and caused them to be especially vigilant in the presence of the discriminative stimulus for lying. Finally, it is difficult for me to let the Maslach experiment pass in a polite fashion. To introduce a procedural artifact in one's own experiment is one thing. To attempt to pawn that artifact off on someone else's experiment and then announce it as the discovery of an alternative explanation of that researcher's findings is quite another. To proclaim this series of mistakes in a major journal makes a nuisance of such foolishness for one's colleagues.

experiment, could have a similar effect. The problem of subtle procedural requirements which are essential for a successful "replication", but not disclosed in the published record of a researcher's methods, is not uncommon in psychological research. Physiological investigations of memory transfer via brain extracts is a popular example,

The second experiment offers little in the way of direct accomplishments. In that experiment an attempt was made to extend the credibility cueing effect beyond the impersonal light cueing situation employed in the first experiment to a situation which involved interpersonal cues. Had the effect occurred with the use of videotaped interviewers as discriminative stimuli for truth and lie telling, this would have broadened the generalizability of the credibility cueing procedure and thereby increased the external validity of the results it can produce. Instead, this experiment produced a partial credibility cueing effect which did not lend itself to unequivocal interpretation. Recall accuracy for trigram operations was higher in the true statement-truth interviewer condition than in any of the remaining three conditions which combined statement and interviewer cues. While this result was regarded as a statement cueing effect for the truth interviewer (cues previously associated with truth telling induced belief in verbal statements made in their presence), interpretation was complicated by a procedural oversight which allowed the interviewer's role (truth or lie cue) to become confounded with his identity (interviewer A or interviewer B). Because one of the two interviewers served as a truth cue for almost all of the subjects, the possibility arose that the

statement cueing effect was not simply the product of his role as a truth cue, but was to some extent peculiar to his identity. Presumably this possible interpersonal cueing effect would have been mediated by the interviewer's appearance, speech, manner, etc. Likewise, the absence of a statement cueing effect for the lie interviewer could have stemmed from some interpersonal aspect of the interviewer who served as a lie cue for most of the subjects. Rather than answering a question of generalizability, the results of the second experiment raised a question as to whether some stimulus persons are more likely to produce credibility cueing effects than others.

By factorially varying the role of live credibility cues (truth or lie) with their identity (interviewer A or interviewer B), the third experiment examined this question. In addition, this experiment retested the internal cue distinctiveness hypothesis. The results suggested that when two persons serve as discriminative stimuli for truth and lie telling, they can have a differential capacity for producing credibility cueing effects. The interactive effect of external cues predicted by the self-perception theory was obtained by one interviewer but not the other. While the manipulation of internal cue distinctiveness was unsuccessful for subjects who attempted to recall in the presence of the interviewer who obtained the effect, a modification of the distinctiveness concept was invoked to explain the differential credibility cueing effects. It was assumed that the rapid pace imposed on recall by the interviewer who obtained the effect interfered with subjects' retrieval of word operations, thus rendering them as inaccessible as trigram operations.

This did not happen to subjects who attempted to recall in the presence of the other interviewer. Hence, the intended manipulation of two levels of internal cue distinctiveness within the interviewer identity variable was unsuccessful. Instead, the discrepancy in pace imposed on recall by the two interviewers inadvertently resulted in a manipulation of distinctiveness between the identity variable. In effect, internal cues appeared to have been rendered less distinct in the presence of the interviewer who obtained the credibility cueing effect, an interpretation which is consistent with the self-perception theory. It was suggested that the other interviewer's failure to obtain the effect for recall of trigram operations was due to the transmission of "personality" cues. Although a single anomaly in the recall data produced in his presence could not be accounted for, a minimal interpretation is that it provides evidence of the cueing procedure's lack of robustness.

This third experiment accomplishes two things: it successfully extends the cueing procedure to a situation involving interpersonal credibility cues, and indicates that some stimulus persons are more likely to produce credibility cueing effects than others. This clearly increases the generalizability of Bem's procedure and the external validity of results produced by it. Because the explanation offered for the latter finding is somewhat speculative, two lines of follow-up research are suggested. One concerns the investigation of retrieval-based interventions designed to optimize the effects of external behavioral and situational cues in the self-perception of subjective states. The other concerns the exploration of interpersonal bases for differential

credibility cueing effects, including those suggested by attribution theoretic sources such as the correspondent inference theory.

Bibliography

- Abelson, R. P. Simulation of social behavior. In G. Lindzey and E. Aronson (Eds.), Handbook of social psychology. Vol. 2 (2nd ed.). Reading, Mass.: Addison-Wesley, 1968.
- Abelson, R. P., and Rosenberg, M. J. Symbolic psycho-logic: A model of attitudinal cognition. Behavioral Science, 1958, 3, 1-13.
- Anderson, A. R., and Moore, O. K. Models and explanations in the behavioral sciences. In G. J. DiRenzo (Ed.), Concepts, theory and explanation in the behavioral sciences. New York: Random House, 1966.
- Archer, E. G. A re-evaluation of the meaningfulness of all possible CUC trigrams. Psychological Monographs, 1960, 74, No. 497.
- Aronson, E. The rationalizing animal. Psychology Today, 1973, 6, 46-52.
- Arrowood, A. J., and Ross, L. Anticipated effort and subjective probability. Journal of Personality and Social Psychology, 1966, 4, 57-64.
- Arrowood, A. J., Wood, L., and Ross, L. Dissonance, self-perception and the perception of others: A study in cognitive cognitive dissonance. Journal of Experimental Social Psychology, 1970, 6, 304-315.
- Aubrey, A. S., and Caputo, R. R. Criminal interrogation. Springfield, Ill.: C. C. Thomas, 1965.
- Bandler, R. J., Madaras, G. R., and Bem, D. J. Self-observation as a source of pain perception. Journal of Personality and Social Psychology, 1968, 9, 205-209.
- Bem, D. J. An experimental analysis of beliefs and attitudes. (Doctoral dissertation, University of Michigan). Ann Arbor, Mich.: University Microfilms, 1964, No. 64-12.

- Bem, D. J. An experimental analysis of self-persuasion. Journal of Experimental Social Psychology, 1965, 1, 199-218.
- Bem, D. J. Inducing belief in false confessions. Journal of Personality and Social Psychology, 1966, 3, 707-710.
- Bem, D. J. Self-perception: An alternative interpretation of cognitive dissonance phenomena. Psychological Review, 1967, 74, 183-200 (a).
- Bem, D. J. Reply to Judson Mills. Psychological Review, 1967, 74, 536-537 (b).
- Bem, D. J. The epistemological status of interpersonal simulations: A reply to Jones, Linder, Kiesler, Zanna, and Brehm. Journal of Experimental Social Psychology, 1968, 4, 270-274.
- Bem, D. J. Beliefs, attitudes, and human affairs. Belmont, Calif.: Wadsworth, 1970.
- Bem, D. J. Self-perception theory. In L. Berkowitz (Ed.), Advances in experimental social psychology. Vol. 6. New York: Academic Press, 1972. Pp. 1-62.
- Bem, D. J., and McConnell, H. K. Testing the self-perception explanation of dissonance phenomena: On the salience of premanipulation attitudes. Journal of Personality and Social Psychology, 1970, 14, 23-31.
- Brehm, J. W., and Cohen, A. R. A re-evaluation of choice alternatives as a function of their number and qualitative similarity. Journal of Abnormal and Social Psychology, 1959, 58, 373-378.
- Brehm, J. W., and Cohen, A. R. Explorations in cognitive dissonance. New York: Wiley, 1962.
- Castaneda, C. The Teachings of Don Juan: A Yaqui Way of Knowledge. New York: Ballantine Books, Inc., 1968.
- Castaneda, C. A Separate Reality: Further Conversations with Don Juan. New York: Simon and Schuster, 1971.

- Castaneda, C. Journey to Ixtlan: The Lessons of Don Juan. New York: Simon and Schuster, 1972.
- Chapanis, N. P., and Chapanis, A. Cognitive dissonance: Five years later. Psychological Bulletin, 1964, 61, 1-22.
- Chappell, V. C. (Ed.) The philosophy of mind. Englewood Cliffs, N.J.: Prentice-Hall, 1962.
- Dutton, D. G., and Douglas, R. L. Cognitive cognitive dissonance revisited: a further look at the dissonance - self-perception controversy. Canadian Journal of Behavioral Science, 1972, 4, 64-74.
- Elms, A. C. Social Psychology and Social Relevance. University of California, Davis: Little, Brown and Company, 1972.
- Elms, A. C., and Janis, I. L. Counterattitudinal attitudes induced by consonant versus dissonant role playing. Journal of Experimental Research in Personality, 1965, 1, 50-60.
- Festinger, L. A theory of cognitive dissonance. Stanford: Stanford University Press, 1957.
- Festinger, L., and Carlsmith, J. M. Cognitive consequences of forced compliance. Journal of Abnormal and Social Psychology, 1959, 58, 203-210.
- Freedman, J. L. Role playing: Psychology by consensus. Journal of Personality and Social Psychology, 1969, 13, 107-114.
- Harré, R., and Secord, P. F. The explanation of social behavior. Oxford: Basil Blackwell and Mott, 1972.
- Hayes, W. L. Statistics for psychologists. New York: Holt, Rinehart and Winston, 1964.
- Heider, F. Attitudes and cognitive organizations. Journal of Psychology, 1946, 21, 107-112.

- Heider, F. The psychology of interpersonal relations. New York: Wiley, 1958.
- Inbau, F. E., and Reid, J. E. Lie detection and criminal interrogation. Baltimore, Maryland: Williams and Wilkins, 1967.
- Jones, E. E., and Davis, K. E. From acts to dispositions. In L. Berkowitz (Ed.), Advances in experimental social psychology. Vol. 2. New York: Academic Press, 1965. Pp. 219-266.
- Jones, E. E., Davis, K. E., and Gergen, K. Role playing variations and their informational value for person perception. Journal of Abnormal and Social Psychology, 1961, 63, 302-310.
- Jones, E. E., and Harris, V. A. The attribution of attitudes. Journal of Experimental Social Psychology, 1967, 3, 1-24.
- Jones, E. E., and Nisbett, R. E. The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, and B. Weiner (Eds.), Attribution: Perceiving the causes of behavior. New York: General Learning Press, 1972.
- Jones, R. A., Linder, D. E., Kiesler, C. A., Zanna, M., and Brehm, J. W. Internal states or external stimuli: Observers' attitude judgments and the dissonance theory - self-persuasion controversy. Journal of Experimental Social Psychology, 1968, 4, 247-269.
- Jones, R. G. Forced compliance dissonance predictions: obvious, non-obvious, or non-sense? Paper presented at the meeting of the American Psychological Association, New York, September, 1966.
- Jordan, N. The mythology of the non-obvious - autism or fact? Contemporary Psychology, 1964, 9, 140-142.
- Kelley, H. H. Attribution theory in social psychology. In D. Levine (Ed.), Nebraska symposium on motivation. Vol. 15. Lincoln: University of Nebraska Press, 1967. Pp. 192-238.

- Kintsch, W. Learning, memory, and conceptual processes. New York: Wiley, 1970.
- Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, Calif.: Wadsworth, 1970.
- Kuhn, T. S. The structure of scientific revolutions. Chicago: University of Chicago Press, 1962.
- Linder, D. E., and Jones, R. A. Discriminative stimuli as determinants of consonance and dissonance. Journal of Experimental Social Psychology, 1969, 5, 467-482.
- Little, B. R. Psychological man as scientist, humanist, and specialist. Journal of Experimental Research in Personality, 1972, 6, 95-118.
- Maslach, C. The "truth" about false confessions. Journal of Personality and Social Psychology, 1971, 20, 141-146.
- McGuire, W. J. Inducing resistance to persuasion. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology. Vol. 1. New York: Academic Press, 1964, 192-229.
- McGuire, W. J. Attitudes and opinions. Annual Review of Psychology, 1966, 17, 475-514.
- McGuire, W. J. Some impending reorientations in social psychology: some thoughts provoked by Kenneth Ring. Journal of Experimental Social Psychology, 1967, 3, 124-139.
- McGuire, W. J. The nature of attitudes and attitude change. In G. Lindzey and E. Aronson (Eds.), Handbook of Social Psychology. Reading, Massachusetts: Addison, Wesley, 1969.
- McGuire, W. J. The yin and yang of progress in social psychology: Seven Koan. Journal of Personality and Social Psychology, 1973, 26, 446-456.

- Mead, G. H. Mind, self, and society. Chicago: University of Chicago Press, 1934.
- Mills, J. Comment on Bem's "Self-perception: An alternative interpretation of cognitive dissonance phenomena." Psychological Review, 1967, 74, 535.
- Newcomb, T. M. An approach to the study of communicative acts. Psychological Review, 1953, 60, 393-404.
- O'Hara, C. E. Fundamentals of criminal investigation. Springfield, Ill.: C. C. Thomas, 1970.
- Osgood, C. E., and Tannenbaum, P. H. The principle of congruity in the prediction of attitude change. Psychological Review, 1955, 62, 42-55.
- Penner, D. D., and Patten, R. L. Comment to "The epistemological status of interpersonal simulations: A reply to Jones, Linder, Kiesler, Zanna, and Brehm." Representative Research in Social Psychology, 1970, 1, 62-66.
- Piliavin, J. A., Piliavin, I. M., Loewenton, E. P., McCauley, C., and Hammond, P. On observers' reproductions of dissonance effects: The right answers for the wrong reasons? Journal of Personality and Social Psychology, 1969, 13, 98-106.
- Ring, K. Experimental social psychology: some sober questions about some frivolous values. Journal of Experimental Social Psychology, 1967, 3, 113-123.
- Rosenberg, M. J. An analysis of affective-cognitive consistency. In C. I. Holland and M. J. Rosenberg (Eds.), Attitude organization and change. New Haven: Yale University Press, 1960.
- Rosenberg, M. J. When dissonance fails: on eliminating evaluation apprehension from attitude measurement. Journal of Personality and Social Psychology, 1965, 1, 28-42.

- Ryle, G. The concept of mind. London: Hutchinson, 1949.
- Schachter, S. The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.), Advances in experimental social psychology. Vol. 1. New York: Academic Press, 1964. Pp. 49-80.
- Schachter, S., and Singer, J. E. Cognitive, social, and physiological determinants of emotional state. Psychological Review, 1962, 69, 379-399.
- Scriven, M. A study of radical behaviorism. In H. Feigl and M. Scriven (Eds.), Minnesota studies in philosophy of science. Vol. 1. Foundations of Science and the concepts of psychology and psychoanalysis. University of Minnesota Press, 1956. Pp. 88-131.
- Shapiro, F. C. Whitmore. London: McMillan, 1970.
- Skinner, B. F. Verbal behavior. New York: Appleton-Century-Crofts, 1957.
- Stratton, G. M. Vision without inversion of the retinal image. Psychological Review, 1897, 4, 341-360.
- Swets, J. A. Signal detection and recognition by human observers. New York: Wiley, 1964.
- Valins, S. Cognitive effects of false heart-rate feedback. Journal of Personality and Social Psychology, 1966, 4, 400-408.
- Valins, S., and Nisbett, R. E. Attribution processes in the development and treatment of emotional disorders. In E. E. Jones, D. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, and B. Weiner (Eds.), Attribution: Perceiving the causes of behavior. New York: General Learning Press, 1972.
- Valins, S., and Ray, A. A. Effects of cognitive desensitization on avoidance behavior. Journal of Personality and Social Psychology, 1967, 7, 345-350.
- Winett, R. A. Attribution of attitude and behavior change and its relevance to behavior therapy. Psychological Record, 1970, 20, 17-32.

- Yaryan, R. B., and Festinger, L. Preparatory action and belief in the probable occurrence of future events. Journal of Abnormal and Social Psychology, 1961, 63, 603-606.
- Zimbardo, P. G., Cohen, A., Weisenberg, M., Dworkin, L., and Firestone, I. The control of experimental pain. In P. G. Zimbardo (Ed.), The cognitive control of motivation. Glenview, Ill.: Scott, Foresman, 1969. Pp. 100-125.