THE SUBJECT'S HYPOTHESIS: ITS DETERMINANTS
AND ITS EFFECT ON RESEARCH DATA

by
RALPH KIRK SAFFORD
B.A., Beloit College, 1964

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS

in the Department
of
Psychology

We accept this thesis as conforming to the required
standard

THE UNIVERSITY OF BRITISH COLUMBIA
January, 1971
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 2/8/71
ABSTRACT

Subjects' hypotheses about the purposes of experiments, regardless of the accuracy of such hypotheses, may contaminate research data. Experiment I was designed to assess how readily subjects generate hypotheses about experiments and what effect such hypotheses have on performance. In the context of a personality impressions experiment, forty subjects participated either in one of two bogus hypothesis conditions, in which they were given unauthorized information about the purpose of the experiment by an accomplice posing as a subject, or in a no bogus hypothesis condition, in which they received no unauthorized information but were interviewed at the experiment's termination for self-generated hypotheses. Bogus hypotheses were not found to have affected subjects' task performance and only two of the fourteen subjects interviewed reported an attempt to generate an hypothesis. This latter result was interpreted as contradicting the notion that subjects are strongly motivated to figure out the purpose of an experiment. It was hypothesized that subjects are indifferent towards the purposes of research generally, and that certain types of experimental stimuli must be present in an experiment in order to arouse subjects to speculate about research purposes. Experiment II, designed to test this hypothesis, investigated the speculation arousal function of two such types of stimuli - experimental rationales and sensitization tasks. Thirty-six subjects participated in one of four conditions provided by
orthogonal manipulation of the two treatment variables and had their level of suspicion, apprehension, and speculation about the experiment assessed by a brief post-experimental questionnaire. Neither factor was shown to have an arousal function as no significant differences were obtained. Other factors possibly involved in arousing subject speculation were discussed.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>Experiment I:</td>
<td>7, 8, 15, 22</td>
</tr>
<tr>
<td></td>
<td>- Method</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Results</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Discussion</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>Experiment II:</td>
<td>29, 31, 34, 39</td>
</tr>
<tr>
<td></td>
<td>- Method</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Results</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Discussion</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Footnotes</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>Appendices</td>
<td>47</td>
</tr>
</tbody>
</table>
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Table I.</th>
<th>Means and standard deviations for hypothesis conditions</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>........................................................................</td>
<td>16</td>
</tr>
<tr>
<td>Table II.</td>
<td>Summary of analysis of variance: Experiment I</td>
<td>............</td>
</tr>
<tr>
<td>Table III.</td>
<td>Means and standard deviations for suspicion, apprehension, and purpose questions</td>
<td>............</td>
</tr>
<tr>
<td>Table IV.</td>
<td>Summary of analysis of variance: Experiment II</td>
<td>............</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

Investigators (Kelman, 1967; Orne, 1962; Riecken, 1962; Schultz, 1969) have argued that subjects who participate in psychological research do not merely passively respond to the experimental stimuli but rather actively attempt to ascertain the true purpose of an experiment. Orne has stated, "The subject's performance in an experiment might almost be conceptualized as problem-solving behavior; that is, at some level he sees it as his task to ascertain the true purpose of the experiment and respond in a manner that will support the hypothesis being tested (1962, p. 779)." Such attempts on the part of subjects to obtain insight into an experiment pose methodological problems for psychologists because they may be a source of artifact. A subject's "hunch" or hypothesis about the real purpose of an experiment may significantly determine his performance in that experiment irrespective of the experimental treatment provided.

Investigation of the subject's hypothesis has in practice meant investigation of only a subset of hypotheses - hypotheses which indicate awareness of the experimental contingencies due to demand characteristics (Page 1966, 1969, in press; Page and Lumia, 1968; Sherman, 1967; Silverman, 1968; Silverman and Regula, 1968). Research on awareness due to demand characteristics is based upon a paper by Orne (1962), in which he proposes that demand characteristics are a source of artifact in psychological experimentation. Demand characteristics are defined as "the totality of cues which convey an experimental hypothesis to the
subject (1962, p. 779)." In demand awareness research, the typical approach is to re-examine some experiment or experimental paradigm the results of which have been "rather uncritically accepted into the psychological body of knowledge (Page 1968, p. 59)," in an attempt to show that the results can be attributed to subjects' awareness of the experimental hypothesis. The procedure is to interview subjects after the experiment and classify them as "aware" or "unaware" based upon their ability to verbalize the experimental hypothesis, and on the basis of this classification determine whether or not the experimental effect can be accounted for by aware subjects alone. (This procedure has been followed in all demand awareness studies except those of Sherman, [1967] and Silverman, [1968] in which no measure of awareness was obtained.)

In all of the demand awareness studies done to date, investigators have found that experimental effects formerly attributed to some treatment variable can be accounted for by subject awareness of demand characteristics. Subjects have been found to be demand-aware in the following experimental paradigms: figure-ground perception (Page, 1968); classical conditioning of attitudes (Page, 1969); communicator credibility (Page, in press); verbal operant conditioning (Page & Lumia, 1968); and the effects of distraction on persuasibility (Silverman & Regula, 1968). Further, attitude change has been shown to be a function of demand characteristics in two studies in which demand characteristics were manipulated (Sherman, 1967; Silverman, 1968). These studies show rather strikingly that demand characteristics are present in a variety of experimental situations and that such cues can affect subjects' experimental performance.
Worthy as research on demand awareness has been, it has unfortunately been the only work done on subjects' hypotheses. The possibility that inaccurate hypotheses (i.e., hypotheses which do not correspond to the experimental hypothesis) may be an important source of artifact has not been investigated. That all subjects' hypotheses - regardless of accuracy - ought to be investigated, is indicated by the following considerations.

First, there is no a priori reason for supposing that an inaccurate hypothesis is any less a potential data contaminant than an accurate hypothesis. All subjects who formulate hypotheses believe with varying degrees of assuredness that their hypotheses are accurate. Since demand awareness research has shown that accurate hypotheses can contaminate data, it is only plausible to assume that inaccurate hypotheses can do likewise.

Secondly, there is reason to believe that subjects are less likely to generate correct hypotheses in experiments than they are to generate incorrect ones. The concentration of research on demand awareness may in part be the result of a belief that subjects can generate accurate hypotheses (i.e., become demand aware) in most of the experiments they encounter. It can be argued, however, that subjects are not likely to be able to generate accurate hypotheses very often. For demand awareness to be considered the rule in psychological research it must be true that demand characteristics are present in most experiments and that subjects are perceptive to the extent of being able to distinguish demand characteristics from irrelevant stimuli. There is little reason to suppose that demand characteristics are present in most experiments. Demand characteristics,
being cues which convey to the subject the experimenter's intentions, are precisely those cues which experimenters characteristically take care to eliminate from their designs. Experimenters must succeed in this enterprise at least occasionally. Indeed, it seems to the author that the most reasonable position to take on this matter is that demand characteristics are present in some, but not most, studies, their presence being primarily a function of the care taken by the experimenter to eliminate such cues. Demand awareness research to date provides evidence to support the notion that demand characteristics are present in some studies, but it cannot be construed as evidence that demand characteristics are present in most studies. When demand characteristics are present in a study (and even if they were present in all studies), the generation of accurate hypotheses is still not assured. Demand characteristics have the potential for conveying to the subject the experimenter's hypothesis; other experimental events just as surely have the potential of misleading the subject about the experimenter's hypothesis. Subjects have no way of knowing which are the "right" cues. For these reasons then, the generation of accurate hypotheses appears improbable.

It can be argued on the other hand that erroneous hypotheses are more common in experiments. Orne (1962), Kelman (1967), Riecken (1962), and Schultz (1969) all for various reasons state that subjects are highly motivated to ascertain the true purposes of research. Desire for awareness does not necessarily, as the demand characteristics advocates would have it, insure awareness. What it may insure, however, is that subjects will have hypotheses. That most of these hypotheses are incorrect is not important; it is only important that
subjects think them to be correct. Thus it appears plausible that psychological research is contaminated not so often by subjects who are aware of an experiment's purpose as by subjects who think (mistakenly) that they are aware. Because inaccurate hypotheses are just as likely to contaminate research data as accurate hypotheses and because inaccurate hypotheses may occur more often in experiments, research in this area should not be restricted to accurate hypotheses alone. All the hypotheses of subjects deserve investigation.

If the subject's hypothesis is to be considered a data contaminant of some significance it must be shown, first, that subjects will readily engage in speculation about an experiment's purpose in a variety of experimental situations (indicating that motivation to speculate is strong and remains so more or less independently of the type of experiment), and second, that such speculation affects subjects' experimental performance. The demand awareness literature contains ample evidence that a "correct" hypothesis can affect performance. It remains to be determined whether or not an incorrect hypothesis can do likewise. By studying the performance consequences of several hypotheses generated about the same experiment, this determination can perhaps be made.

In a preliminary study on the subject's hypothesis, the author attempted to determine the nature and extent of subject speculation in a given experimental setting as well as its consequences for subject experimental performance. In the context of a pretest-posttest opinion change experiment, some subjects were surreptitiously given information ("hypotheses") about what type of
change was expected and some subjects, given no information, were queried about spontaneously generated hypotheses by means of an extensive post-experimental questionnaire. The resultant opinion change for subjects in the hypothesis manipulation conditions was in the predicted direction but the difference was not statistically significant. The amount of subject speculation in the no information condition was quite high — approximately 90% of those subjects had some ideas concerning the purpose of the experiment. Not only was there some variety in the hypotheses generated, but more than a few of them were based upon experimental events that had no demand characteristics function. Some subjects simply attended to the "wrong" cues. It seems then that even in experiments the purpose of which ought to be fairly obvious to subjects, it is possible to misread the signs and to generate an inaccurate hypothesis.

As a result of these findings the need for further studies was indicated. The two studies reported below were undertaken for this purpose. The first experiment, designed partially as a replication of the author's preliminary study, attempted to determine whether or not subjects will try to figure out the purpose of an experiment and whether or not hypotheses affect performance. It was predicted that subjects would readily generate hypotheses and that such hypotheses would affect performance. The second experiment was concerned with factors that cause subject speculation. As the second experiment was based upon interpretation of the findings of the first experiment, discussion of its rationale and purposes will be deferred until later.
CHAPTER II
EXPERIMENT I

The purpose of this experiment was to determine a) to what extent subjects generate hypotheses about the purpose of experiments and b) whether or not such hypotheses affect performance. The context in which these two factors, amount and effects of subject speculation, were tested consisted of an attitude change paradigm described to subjects as an experiment on personality impressions. Subjects were instructed to read a written communication in order to obtain an impression of the author's personality and then asked to fill out two forms, one related to their obtained personality impression and one, presented as a "control" measure, related to their opinions about the communication. One group of subjects (No Bogus Hypothesis condition) after completing the experiment were interviewed individually in order to ascertain how many of them generated hypotheses about the purpose of the study. A second and third group of subjects (Bogus Hypothesis condition) were, during the course of the experiment, given unauthorized information about the purpose of the experiment by an accomplice posing as a subject (these bogus hypotheses were designed to simulate hypotheses that subjects might be expected to generate on their own). As the accomplice's information had a direct bearing on the form that was used to measure subject opinion, this form served as the dependent measure in the study.

One fault with the preliminary study was that the design was such that subjects could generate an appropriate hypothesis
relatively easily. As ease of guessing an experiment's purpose is considered to be a factor in influencing only the accuracy and not the quantity of hypotheses, a design in which purposes were more difficult for subjects to guess was used in this study. It was felt that by measuring opinions once only and emphasizing the personality impressions rationale, the experiment's "real" purpose would be rendered relatively obscure.

It was predicted, first, that a considerable number of subjects in the No Bogus Hypothesis condition would report hypotheses about the experiment's purpose (although no precise prediction was made here, it was expected that at least half of the subjects would report hypotheses). It was expected that these hypotheses would be variable in content and based, not upon a few "significant" cues as might be predicted from a demand characteristics viewpoint, but upon a variety of experimental stimuli. Secondly it was predicted that subjects in the two Bogus Hypotheses conditions would express opinions that differed from those expressed by subjects in the No Bogus Hypothesis condition, opinions that would be a function of the bogus hypothesis to which they were exposed.

**METHOD**

**Subjects**

Each of forty volunteer subjects enrolled in introductory psychology courses during summer session at Vancouver City College participated in one of the three conditions of the experiment. The sample consisted of an equal number of males and females.
Subjects were run in small groups ranging in size from three to six persons. Only two subjects reported that they had had prior experimental experience, and no subject reported any foreknowledge of this experiment.

Procedure

The experimental procedure outlined below was the same for all subjects participating in the experiment.

At the start of the experiment subjects were given a "cover story" about the experiment's purpose and procedures. Subjects were told that the experiment was concerned with how "personality impressions are formed." They were told that they would be given a written article to read from which they were to attempt to obtain an impression of the author's personality, and that their impressions would be assessed on a Personality Impressions Scale. Subjects were also informed that they would be given an additional written task the purpose of which was to "control factors which may have unintentionally influenced" their obtained personality impression.

Subjects then read a 550 word communication (Appendix A) arguing against the idea that soon, college graduates would have to undertake advanced graduate training to obtain employment positions open to them today. This communication was previously used by Papageorgis (1970).

After subjects read the communication they completed the Personality Impressions Scale (Appendix B). Subjects had three tasks to perform on it: respond to twelve seven-interval scales with adjectival opposites descriptive of personality traits at the
endpoints (e.g. "Assertive-Submissive," "Intelligent-Unintelligent," etc.); rate the strength of the personality impression they obtained from reading the communication; and write a three-sentence description of the author. The Scale itself was preceded by a set of written instructions.

Next a five-item Opinionnaire (Appendix C), described to subjects as a control task, was administered. The Opinionnaire consisted of five statements about the communication, each statement followed by a seven interval scale with endpoints labelled "agree-disagree." The five statements were as follows: "The article, on the whole, was a good one"; "The evidence presented by the author was objective and factual"; "The arguments used by the author were designed to play on the reader's emotions"; "The belief that advanced graduate study will soon be essential for employment is false"; and, "The author of the article was sincere."

At the end of the experiment all subjects were asked not to discuss the experiment with fellow students.

Treatment conditions

Subjects participating in the Bogus Hypothesis (BH) conditions received (to them) apparently unauthorized information about the purpose of the experiment from a female accomplice posing as a subject. Half of the subjects received positive information about the experimenter's intentions (Pro Bogus Hypothesis), and half received negative information (Con Bogus Hypothesis). The accomplice made her "pitch" just prior to the administration of the Opinionnaire while the experimenter was absent from the experimental room on the
pretext that he had forgotten to bring the Opinionnaire forms with him. The accomplice was instructed to impart the following information to the respective groups:

PRO BOGUS HYPOTHESIS

I have a friend who was in this experiment before and she told me what it's all about. This guy isn't testing personality. He's really testing our opinions about the article. Apparently he wants to see how open minded students are, see if they're willing to be convinced by new ideas, like in the article. That's why he's going to test our opinions.

CON BOGUS HYPOTHESIS

I have a friend ... opinions about the article. Apparently he wants to see how gullible students are, see how easily they can be made to believe a foolish idea, like in the article. That's why he's going to test our opinions.

The accomplice's information was designed to sensitize subjects to an alternative interpretation of the purpose of the experiment and to simulate the sort of hypothesis that a subject might generate if left to his own devices - that persuasion was the real purpose of the study. The experimenter returned to the experimental room after a two minute absence and administered the Opinionnaire.

Subjects in the No Bogus Hypothesis (NBH) condition were given no unauthorized information about the purpose of the experiment and were interviewed individually afterwards to determine what hypotheses, if any, they had generated on their own during the course of the experiment. At the end of the experiment proper, subjects were requested to remain for an interview. The experimenter then told subjects that his interest in conducting the experiment was not to study personality impressions but to find out what their personal reactions to such an experiment would be. They were told that the
sole function of the experiment was to give them experimental experience for the forthcoming interview. The experimenter then discussed (without revealing his interest in hypotheses) the importance of studying subjects' "reactions to experiments" and asked subjects to return to the experimental room at ten minute intervals for the interviews. The purpose of the interview was explained to subjects in this way so that obstacles to frank discussion might be eliminated and the possibility of an experimenter-subject "pact of ignorance" (Orne, 1959a) avoided.

Each interview took about ten minutes and was tape recorded. All subjects in the No Bogus Hypothesis condition were interviewed, and no subject had to wait more than thirty minutes for his interview session. During the interview, subjects were questioned about their prior attitudes toward and knowledge of experiments, their reactions to each of the various tasks in this experiment, and specifically about any suspicions or hypotheses they had about the experiment's real purpose. If a subject reported an hypothesis, he was asked when he had formulated the hypothesis and what had caused him to formulate it. The questioning was designed to be thorough and the interviewer attempted to follow up any promising leads with further questions.

Analysis

The Opinionnaire responses of subjects in both Pro and Con BH conditions were summed across statements for each subject and the summed scores were compared with the summed Opinionnaire scores of subjects in the MBH condition. To facilitate analysis, scoring for statement #3 was reversed. It was predicted that, in line with their
respective bogus hypotheses, subjects in the Pro Hypothesis group would agree more and subjects in the Con Hypothesis group would agree less with Opinionnaire statements, than subjects in the No Bogus Hypothesis condition.

A considerable number of subjects in the HBH condition were expected to report hypotheses. A subject was not classified as having an hypothesis unless all four of the following criteria were met: the subject must report that he had attempted to figure out the experiment's purpose (thus indicating that he had been sceptical of the experiment's stated purpose); he must be able to report an hypothesis that was specific enough to affect his experimental behavior (vague hunches were not acceptable); his belief in his hypothesis must not be tentative; and he must have arrived at his hypothesis prior to the experiment's termination. Speculation that fails to meet these criteria will not, theoretically at least, have the power to contaminate performance, and were not considered hypotheses.

Two problems were encountered in the study - one problem having to do with implanting the hypothesis, the other having to do with the Opinionnaire.

During one of the sessions of the Con Hypothesis condition, while the accomplice was in the act of telling subjects what the experiment was about, a subject questioned her authenticity, implying that she was a "plant." Devastating as this accusation ought to have been for the accomplice's ability to maintain verisimilitude, she reported that she was able to manage the situation so that, in the end, the
subject was himself discredited rather than the accomplice. (He was apparently playing a hunch and was easily dissuaded.) A welcome assist was even provided the accomplice by several subjects who ridiculed the "plant" notion. Since all subjects, including even the one who spoke out, appeared to side with the accomplice in the end, the data from this session were not excluded from the analysis. No such problems were encountered in the other bogus hypothesis sessions and for them all indications are that the accomplice's words went unquestioned (for example, following the accomplice's disclosure a subject in one session said, "Yeah, that's the way they [psychologists] do things").

The second difficulty had to do with one statement on the Opinionnaire (h). After the experiment was completed it was discovered that this statement ("The belief that advanced graduate study will soon be essential for employment is false.") may have been a source of confusion to subjects. The statement was worded rather awkwardly ("The belief that advanced graduate study will be essential ... is false" is an awkward way of saying "Advanced graduate study will not be essential ..."), and therefore some subjects may have inadvertently indicated an attitude toward the statement that was opposite from the one they actually held. To check on this possibility, the statement was tested for clarity on ten subjects not associated with the original experiment. Six subjects were confused by the statement. Since subjects in the experiment may have been similarly confused, the data from this statement were not used in the analysis.
RESULTS

Opinionnaire Data

The Opinionnaire data for the hypothesis comparison are summarized in Table I (one subject from the iNH condition was randomly eliminated to provide equal-sized groups for the analysis). The higher the mean score, the greater the agreement with the Opinionnaire statement(s). It was predicted that Opinionnaire scores for the Pro and Con BH conditions would be higher and lower respectively than Opinionnaire scores for the iNH condition. Note that the overall means bear out only half of this prediction: the Con Hypothesis mean ($\bar{x} = 3.75$) is lower than the iNH mean ($\bar{x} = 4.91$), but the Pro Hypothesis mean ($\bar{x} = 3.87$) is not higher than the iNH mean. An analysis of variance performed on the Opinionnaire data (Table II) showed that these means did not differ at conventional levels of significance ($F = 2.62$, $p < .10$). Thus the bogus hypothesis manipulations did not have the predicted differential effect on Opinionnaire performance.

Although the bogus hypothesis manipulation did not have the predicted effect on Opinionnaire performance, the manipulations nonetheless may have had some other effect. Not only are both BH means lower than the iNH mean, but the means for individual Opinionnaire statements in both Pro and Con BH groups are consistently lower than the corresponding means for the iNH condition, sometimes strikingly so (see particularly statement 3/3). Both Pro and Con Bogus hypotheses seem to have affected performance (that is, if the borderline significance is taken to reflect a true difference);
contrary to expectations however, the effect of both bogus hypotheses was the same.

TABLE I
MEANS AND STANDARD DEVIATIONS
FOR HYPOTHESIS CONDITIONS

<table>
<thead>
<tr>
<th>Opinionnaire Statement</th>
<th>Bogus Hypothesis Condition</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Con Hypothesis</td>
<td>No Hypothesis</td>
</tr>
<tr>
<td>1</td>
<td>4.08</td>
<td>4.23</td>
</tr>
<tr>
<td>2</td>
<td>3.31</td>
<td>4.15</td>
</tr>
<tr>
<td>3</td>
<td>3.31</td>
<td>5.62</td>
</tr>
<tr>
<td>5</td>
<td>4.31</td>
<td>5.77</td>
</tr>
<tr>
<td>X</td>
<td>3.75</td>
<td>4.94</td>
</tr>
<tr>
<td>SD</td>
<td>4.81</td>
<td>6.04</td>
</tr>
<tr>
<td>n^d</td>
<td>13</td>
<td>13</td>
</tr>
</tbody>
</table>
**TABLE II**

**SUMMARY OF ANALYSIS OF VARIANCE:**

**EXPERIMENT I**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypothesis</td>
<td>2</td>
<td>89.95</td>
<td>2.62*</td>
</tr>
<tr>
<td>Error</td>
<td>36</td>
<td>34.32</td>
<td></td>
</tr>
</tbody>
</table>

* * p < .10

**Interview Data**

Fourteen subjects participating in the No Bogus Hypothesis condition were interviewed for their hypotheses (the subject excluded from the Opinionnaire data analysis was not excluded here). In all but a few cases, the indications are that subjects responded to the interviewer's questions honestly. For example, the reaction of subjects to questions about their suspicions and hunches was in many cases that of surprise - a difficult response to fabricate effectively. Also, subjects reporting experimental behavior "out of the ordinary" did not display any apparent reluctance to disclose that behavior to the interviewer. Those few subjects whose interview responses were "guarded" were noted by the interviewer, and their reports were given little credence.
Contrary to expectations, no subject had an hypothesis about the purpose of the experiment that measured up to the criteria. One subject came to the experiment with an elaborate hypothesis, but rejected it as unsuitable when faced with experimental events. This subject (#36) thought he was going to be given an intelligence test. When queried further, he said that he thought the purpose of the experiment was a comparison of the IQ of students at his school with the IQ of students at a nearby university. How he arrived at this hypothesis could not be determined, but it apparently arose out of a knowledge, and perhaps apprehension, of deception experiments. He realized part way through the experiment, however, that his hypothesis was not tenable, and rejected it: "After having finished the personality impressions test I really didn't see how it could be an intelligence test and I really didn't see what else you'd be testing."

It is interesting that this subject's belief in his hypothesis - unsuitable though it was - remained unshaken until after he had completed the Personality Impressions Scale. He obviously did not believe that the experimenter's introductory remarks about the experiment's purpose and procedure, and required a wealth of disconfirming evidence (communication, Personality Impressions Scale instructions, etc.) before his belief in his hunch was shaken. When an hypothesis can be maintained in the face of such disconfirming evidence, belief in that hypothesis must be strong. This suggests that subjects whose hypotheses do not conflict with experimental events may place a considerable amount of faith in their hypotheses.
correctness.

Subject #36 displayed all the attributes of a subject with an hypothesis except one: he failed to maintain belief in his hypothesis. Therefore he could not be classified as having an hypothesis by the criteria set forth in this study. The experimental performance of a subject who has and then rejects an hypothesis is not likely to be contaminated by that hypothesis.

Two subjects (#'s 20 and 25) reported that during the course of the experiment they thought it might be about something other than personality impressions. However their hunches were both vague and transitory in nature. Not only could they not say what that "something else" might be, but the thought that the experiment was about something else apparently just "crossed their minds." Aside from such momentary doubts they seemed to maintain belief in the stated purpose of the experiment. These subjects therefore were not classified as having an hypothesis.

None of the other eleven subjects reported anything that could be even remotely considered an hypothesis. No subject interviewed in this experiment therefore, was classified as having an hypothesis.

Not only were hypotheses apparently absent from subjects in this experiment, but suspicion of the experimenter's intentions, a factor which ought to lead subjects to entertain hypotheses, also appeared to be quite low. Only one subject was considered to be suspicious. This subject (#30) had apparently been forewarned about deception experiments by her psychology instructor, and heeded his warning. The relevant part of the interview follows:
Interviewer: When I gave you the instructions at the beginning, did you believe them?

Subject: No. Our psychology professor said that you never know what somebody is testing ... like you may be asked to do a certain task and they can be testing something else.

Interviewer: So when I gave you the instructions you were a little suspicious?

Subject: No. I don't mind. [But a moment later:] I was wondering what your whole test was based on and what you were trying to find.

Although this subject was suspicious, she was unable to come up with an hypothesis about the experiment's purpose: "I tried to figure it out, but didn't have much luck." A little more psychological and experimental sophistication perhaps, and this subject might have been able to figure it out.

One other subject (#22) reported that she was suspicious in the experiment, but her words are suspect. Not only were her interview responses guarded, but she seemed to be making an effort to give only those responses that the interviewer would want to hear. Because her responses to other interview questions did not appear to be forthright, the truthfulness of her reported suspiciousness was cast in doubt. Therefore she was not classified as suspicious.

An interesting interview finding related to subject suspicion and hypothesis testing was the number of subjects reporting awareness of experimental deceptive practices who were not suspicious. Six of the fourteen subjects interviewed reported that they knew about deception experiments (how much they knew about them is another matter; none were truly sophisticated in their knowledge). Of these, only two reported that they suspected that the experimenter was testing
something other than personality impressions (subjects #30 and #36 previously discussed). The other four subjects said that, despite their knowledge of deception, they were not suspicious. These subjects gave one of two reasons for their lack of suspicion. One reason was that they came to the experiment expecting something "horrible" only to find that the experiment was not as they expected, and therefore dismissed the thought of deception from their minds. The other reason given was that to attempt to "second guess" the experimenter was "wrong". It appears then that subject knowledge of experimental deceptive practises, although perhaps a necessary condition for the arousal of suspicion and hypotheses, is not a sufficient condition for their arousal. When an experiment is innocuous or when subjects adopt a "proper" frame of mind, forewarned is not necessarily forearmed.

In conclusion, it is apparent from the interview data that subjects did not generate hypotheses about the purpose of the experiment as expected. Out of fourteen subjects interviewed, no subjects were classified as having an hypothesis and only two subjects were classified as suspicious (assuming that the subject with the aborted "intelligence test" hypothesis ought to be placed in the "suspicious" category). A majority of subjects reported that they accepted the personality impressions rationale completely and that their behavior in the experiment was task-oriented. The remainder of the subjects reported nothing more substantial than vague and fleeting doubts. The occurrence of hypotheses in subjects appears by these data to be rare.
DISCUSSION

Opinionnaire data

The failure of the bogus hypothesis manipulation to produce the predicted effect was surprising, especially since a similar manipulation in the author's preliminary study produced results which, although not significant, were at least in the predicted direction. Two possible explanations can be given for the failure of the manipulation: that the bogus hypotheses were not effectively communicated to subjects; or that not all subjects were desirous of confirming the experimenter's expectations.

If subjects understood that the experimenter was really testing "open mindedness" or "gullibility," some of them may have decided to be negative responders (Masling, 1966). To be negative responders (i.e., to give responses known to be the opposite of what the experimenter desires or, in Masling's terminology, to give responses which "screw" the experimenter), subjects in the Pro Bogus Hypothesis condition would have had to show minimal agreement with statements on the Opinionnaire, and subjects in the Con Bogus Hypothesis condition would have had to show maximal agreement. As subjects in both BH conditions showed (when compared with the mean of subjects in the NBH condition) minimal agreement with the Opinionnaire statements, it appears unlikely that these data are an instance of Masling's "screw you" effect. Subjects in the Con Bogus Hypothesis condition did not give the appropriate "screw you" response. Unless negative responding was used only by subjects in the Pro Bogus Hypothesis condition (possibly as a result of a less plausible rationale in the
"open mindedness" instructions), the Masling interpretation may be ruled out.

It seems much more likely that the results can be accounted for by some sort of communication breakdown. The accomplice did report however that in no case did she have any difficulty in imparting a bogus hypothesis to subjects. The key to what happened therefore may lie in understanding the conditions under which subjects received the accomplice's information.

It was determined from the interviews that most of the subjects in the MBH condition gave credence to the personality impressions rationale provided by the experimenter and were primarily task oriented. Subjects in the BH conditions, prior to receiving the accomplice's information, probably placed a similar amount of faith in that rationale. Discovering that the experimenter was not testing personality impressions probably took them by surprise. Given this situation, it seems intuitively plausible that subjects were more interested in determining the extent of the experimenter's deceit than in learning the experiment's true purpose. Contrast this situation with the one under which subjects received bogus hypotheses in the author's first study. There, subjects reported that they were suspicious from the beginning. They were probably quite interested in learning from the accomplice what the experiment's true purpose was. But in the present study, subjects, since they were probably unsuspecting, may have been caught off guard by the unanticipated deception, and so did not give the accomplice's information the consideration it deserved. The bogus hypothesis manipulation therefore may have been only
effective in arousing distrust and suspicion and not effective in communicating information about the experiment's purpose as intended.

If the above surmise is correct (and it seems to the author a more plausible account of the situation than that some subjects were negative responders), then this experiment must be considered to have inadequately tested the prediction made about the behavioral consequences of having an hypothesis. Further research is necessary.

In light of the possibility that the hypothesis manipulation only succeeded in making subjects suspicious, the Opinionnaire data may be of increased interest. The fact that subjects in both Pro and Con BH groups scored lower (p < .10) on the Opinionnaire than subjects in the NBH condition (who were for the most part not suspicious) suggests that suspiciousness of the experimenter's intent may be a research contaminant in its own right. These findings are in line with those of Stricker, Hessick & Jackson (1967) who in a conformity study found that suspicious subjects conformed less. Further investigation of this "suspicion effect" appears by these data to be recommended.

**Interview findings**

The fact that only 15% of the subjects interviewed in the NBH condition saw reason to doubt the alleged purpose of the experiment may be an important finding despite its unexpectedness. It is at odds with the findings of the author's preliminary study and at odds with the findings of most of the research on demand awareness (e.g. Page, 1968, 1969; Silverman & Regula, 1968), in which a greater percentage of subjects sceptical of an experiment's alleged purpose and aware of
an experiment's true contingencies have been found. Possible causes of this discrepancy merit consideration.

On the surface, it appears that the relative absence of suspicion and speculation in the MBT condition of this experiment may be attributed to the subjects' lack of sophistication in experimental and psychological matters. Only two subjects had participated in a psychology experiment prior to this one, and all were enrolled in introductory psychology courses in which experimental psychology was apparently not heavily emphasized. It is possible that had subjects in this experiment been more sophisticated, they might have been more prone to speculate about its purpose. Evidence from this and other experiments however suggests that subject sophistication may not in fact be of much importance as a factor in determining the amount of suspicion and speculation. First, suspicion has been found in subjects no less naive than subjects in this study. Subjects in the author's preliminary study were just as naïve on the average as subjects in this study, and naïveté did not prevent their being suspicious. Widespread suspicion has even been encountered in a sample of high school students (Stricker, Messick & Jackson, 1967). Secondly, sophistication and experimental experience have not been shown to lead to increased suspicion and speculation. Fillenbaum (1966) exposed subjects to mild deception but found no increase in suspiciousness in a subsequent experiment. Page (1968, 1969) found that subject sophistication was associated with awareness of demand characteristics, but concluded that sophistication was not necessary for awareness because many unsophisticated subjects were aware in his
Holmes (1967) found that prior experimental experience increased the probability of a subject's becoming aware, but also found that more experienced subjects were less inclined to engage in speculation about the purpose of an experiment. The relationship between sophistication and suspicion/speculation is then, not entirely clear. It appears that a sophisticated subject has a better chance of success in his speculations, but it does not appear that he is any more motivated to speculate than is a naïve subject. Naïve subjects have not been found to be any less suspicious than sophisticated subjects. Some other factor appears to underlie suspicion and speculation.

One difference between this study and demand awareness studies is that the experimental paradigms that are chosen for investigation in demand awareness research are chosen because they are believed to contain demand characteristics; the paradigm used in this study was not chosen for that reason. Thus it may be the case that demand characteristics were present in all the paradigms investigated for demand awareness, but were not present in the paradigm investigated in this study. Perhaps this difference accounts for the discrepancy in findings re subject speculation. The presence of demand characteristics in an experiment may determine whether or not subjects will speculate about that experiment. The only function that demand characteristics are supposed to have is to convey to subjects the experiment's purpose, but perhaps they have an arousal function as well: that is, they cause previously unsuspecting subjects to doubt the experimenter's intentions.
If demand characteristics do function to arouse subject scepticism as well as to convey information about an experiment's purpose, the implications concerning the type of artifact being dealt with by investigators in this area are quite important. It has been assumed (Kelman, 1967; Orne, 1962; Reicken, 1962; Schultz, 1969) that the source of the artifact in part resided in subjects' keen interest in ascertaining the true purposes of research. This however may not be the case. What we may have is not a population of overly suspicious subjects, but an indeterminate number of poorly disguised experiments. Subjects may speculate about an experiment only when some procedural irregularity arouses their suspicions. If so, the artifact is controlled not by finding some contaminant-free method of conducting experiments within the context of suspiciousness, but by designing experiments in which demand characteristics and other suspicion arousing cues are eliminated. Martin Orne (1969) has in a recent article outlined procedures that could be used to eliminate such cues from experiments.

If subjects only speculate when aroused and if speculation can be eliminated by appropriate methodology, then determination of the nature and consequences of speculation - i.e., continuation of the research program of this study - may be of only minor importance. Of major importance is determining whether or not certain experimental practises can account for the speculation of subjects heretofore attributed to high motivation. Experiment II therefore was designed as a preliminary investigation of two experimental practises believed
to be instrumental in arousing subject suspicion and speculation. The two practices investigated were the use of sensitization tasks and cover stories.
CHAPTER III
EXPERIMENT II

In the author's preliminary study, most subjects speculated about the experiment's purpose. In the study just reported, speculation was evidently rare. The discrepancy in findings may be accounted for by procedural differences between the two studies. Practises eliminated from Experiment I may have been significant in arousing subjects to speculate in the preliminary study. Two practises which may have been significant in this regard are as follows. In the preliminary study, subjects were not provided a cover story about the purpose of the experiment and had their opinions measured twice. The absence of a cover story and the taking of repeated opinion measures may, either singly or in combination, cause subjects to speculate about an experiment's purpose. These two experimental practises were investigated in this study.

The significance of repeated attitude measures as a factor in arousing subject suspicion and speculation lies in the fact that when a performance measure is administered twice in an experiment, subjects may become sensitized to the possibility that their responses on the two measures will be compared. Thus investigators often make an effort to disguise such measures (or resort to the use of a single measure). If the disguise is somehow inadequate, or if two measures having in fact no relationship are similar-appearing, subjects may become suspicious. Thus any unexplained, similar-appearing measures in an experiment are likely to have an arousal function.
That failure to provide a rationale or cover story for an experiment can result in subject speculation seems plausible. A silent experimenter may be regarded by subjects as one who has "something to hide." Simple curiosity will cause subjects to become interested in finding out what the experimenter is not telling. Even if the experimenter explains his inability to describe the experiment's purpose, subjects will have difficulty functioning without the framework that a rationale would provide. Thus they may find it necessary to make some sense of the course of experimental events.

As sensitization tasks and experiments without rationales seem likely candidates for arousing subject suspicion and speculation, they were investigated in this study. The variables were manipulated orthogonally to provide four treatment conditions (rationale vs. no rationale and sensitization vs. no sensitization). The personality impressions/persuasion experiment format used in the first experiment was also used in this study. The dependent measure was a brief post-experimental questionnaire designed to measure subjects' level of suspicion, apprehension, and speculation. It was predicted that subjects provided with no rationale for the experiment would report greater suspicion, apprehension, and speculation than subjects provided with a rationale, and that subjects receiving the sensitization task would be more suspicious, apprehensive, and speculative than subjects not receiving it.

In addition, several correlational analyses of the data were undertaken. Level of suspicion was correlated with degree of agreement with the communication and extremity of opinion (both as measured by an Opinionnaire) and, in both cases, the correlation was
predicted to be negative. These analyses were conducted because of the arguments presented in the discussion section of the preceding experiment, that suspicion in a subject sample may be a data contaminant in its own right. The prediction that suspicious subjects will agree less with Opinionnaire statements than unsuspicious subjects arises from the results of the previous study, in which the Opinionnaire differences were interpreted as indicating a "suspicion effect." The level of suspicion-extremity of opinion prediction was based on the belief that suspicious subjects are "on guard," and, due to caution, avoid the expression of extreme viewpoints (e.g., avoid marking the endpoints of scales).

METHOD

Thirty-six undergraduate volunteers drawn from introductory psychology courses at the University of British Columbia participated in one of the four treatment conditions. Prior experimental experience of subjects averaged 1.9 experiments (range, 0-6), and, as the experiment was conducted near the end of term, subjects should be considered both experimentally and psychologically sophisticated. The sample was made up of 17 males and 19 females.

The format for this experiment was similar to that used in the first experiment. All subjects, regardless of treatment condition, were given the following tasks to do:

a) read a 550 word persuasive communication about higher education (the same communication used in the first experiment);

b) fill out a Personality Impressions Scale (the same scale as used in the first experiment);
c) fill out an Opinionnaire consisting of five statements about the communication (this form differed from that used in the first experiment only in that the last two statements were re-phrased, as follows: "The conclusions arrived at in the article are essentially correct" and, "The author obviously had ulterior motives for writing the article" - Appendix D);

d) answer a post-experimental Questionnaire (Appendix E) consisting of four questions about their reactions to the experiment. (The questions were as follows: "Were you suspicious of the experimenter's intentions?"; "Were you apprehensive about doing any of the tasks?"; "How much time did you spend attempting to guess the experiment's purpose?"; and a fourth question which was a filler item and was omitted from the analysis.) Each question was followed by a four point scale with intervals labelled as to the amount of the quantity in question - e.g., "not at all suspicious," "slightly suspicious," "moderately suspicious," and "very suspicious."

The two treatment variables, rationale for the experiment and sensitization, were manipulated in the following manner. Subjects participating in the rationale condition received at the start of the experiment instructions providing them with a plausible account of the purposes of the study. They were told that the investigation was being undertaken to determine "how personality impressions are formed." The experimenter told them he was interested in "what type of personality people attribute to an individual given that they have various sorts of information about him," and then went on to explain that they would be given a written article to read from which they were to attempt to obtain an impression of the author's personality.
In addition to these instructions, just prior to the administration of the Opinionnaire, subjects were informed that it was for control purposes, "because [their] opinions about the article may have unintentionally influenced [their] personality ratings." Subjects participating in the no rationale condition received no such account of the experiment. They were told at the beginning simply that they would have to do "various written tasks," and the Opinionnaire was administered without explanation.

In the sensitization task condition, an attitude survey form labelled "Attitudes Toward Higher Education" (Appendix F) was administered prior to the communication. The form consisted of six statements, each statement followed by a five interval scale ranging from "definitely false" to "definitely true." Three of the statements were relevant to some point mentioned in the communication subjects would read later, and three were not relevant. The purpose of the survey form was to make subjects aware of the possibility that it was their opinions about the communication (which would subsequently be measured on the Opinionnaire) and not their impressions of the personality of the author of the communication, that were of interest to the experimenter. Subjects receiving the attitude survey in the rationale condition were told that it was given for control purposes. In the no rationale condition the purpose of the survey was not explained.

In the no sensitization condition the attitude survey was simply not administered. Following the instructions subjects began reading the communication.
A main effect for both variables, rationale and sensitization, was predicted. In addition, a negative correlation was expected to result from the level of suspicion - degree of agreement and level of suspicion - extremity of opinion analyses.

RESULTS

Questionnaire data are summarized in Table III. Note that responses to each question are tabled separately, and that the possible range of scores for each question is from one to four. Across all conditions, the scored responses to each question can be ranked from lowest to highest as follows: apprehension ($\bar{x} = 1.45$), suspicion ($\bar{x} = 1.92$), and guessing about the experiment's purpose ($\bar{x} = 2.14$). All these scores are moderately low with respect to the quantity in question, and the distributions are generally skewed. For example, the overall mean of 1.92 for the suspicion question indicates that on the average, subjects were less than "slightly suspicious" about the experimenter's intentions.

An analysis of variance was performed separately on each question (Table IV). (The data did not appear sufficiently compelling to perform a multivariate analysis of variance on the entire set of data.) A two way analysis of variance performed on data from the suspicion question yielded no significant differences (all $F$s less than 1). The predictions that the absence of a rationale and the presence of a sensitization task create more suspicion were not confirmed.
On the apprehension question, a two way analysis of variance resulted in a difference for the sensitization factor which approached conventional significance levels ($F = 3.76, p \leq .07$). This difference however is in the direction opposite from that predicted: subjects reported more apprehension without sensitization than with it. This clearly does not confirm the hypothesis with respect to the arousal function of sensitization.

### TABLE III
**Means and Standard Deviations**

**For Suspicion, Apprehension, and Purpose Questions**

($N = 9$ per cell)

<table>
<thead>
<tr>
<th>GROUP</th>
<th>RATIONALE</th>
<th></th>
<th></th>
<th>NO RATIONALE</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Suspicion</td>
<td>Apprehension</td>
<td>Purpose</td>
<td>Suspicion</td>
<td>Apprehension</td>
<td>Purpose</td>
</tr>
<tr>
<td>Sensitization</td>
<td>2.11</td>
<td>1.11</td>
<td>2.84</td>
<td>1.89</td>
<td>1.22</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td>(1.15)</td>
<td>(.33)</td>
<td>(.88)</td>
<td>(1.17)</td>
<td>(.44)</td>
<td>(1.09)</td>
</tr>
<tr>
<td>No Sensitization</td>
<td>1.78</td>
<td>1.56</td>
<td>2.33</td>
<td>1.89</td>
<td>1.89</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>(.83)</td>
<td>(1.01)</td>
<td>(.87)</td>
<td>(.05)</td>
<td>(.27)</td>
<td>(1.12)</td>
</tr>
</tbody>
</table>

**Note:** Standard deviations are in parentheses.
### TABLE IV
SUMMARY OF ANALYSIS OF VARIANCE: EXPERIMENT II

#### 1. Suspicion

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitization (A)</td>
<td>1</td>
<td>.25</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Rationale (B)</td>
<td>1</td>
<td>.03</td>
<td>&lt;1</td>
</tr>
<tr>
<td>AXB</td>
<td>1</td>
<td>.25</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Error</td>
<td>32</td>
<td>1.32</td>
<td></td>
</tr>
</tbody>
</table>

#### 2. Apprehension

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitization (A)</td>
<td>1</td>
<td>2.78</td>
<td>3.76*</td>
</tr>
<tr>
<td>Rationale (B)</td>
<td>1</td>
<td>.15</td>
<td>&lt;1</td>
</tr>
<tr>
<td>AXB</td>
<td>1</td>
<td>.10</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Error</td>
<td>32</td>
<td>.74</td>
<td></td>
</tr>
</tbody>
</table>

* p < .07

#### 3. Purpose

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitization (A)</td>
<td>1</td>
<td>.03</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Rationale (B)</td>
<td>1</td>
<td>2.25</td>
<td>2.27</td>
</tr>
<tr>
<td>AXB</td>
<td>1</td>
<td>.25</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Error</td>
<td>32</td>
<td>.99</td>
<td></td>
</tr>
</tbody>
</table>
For the purpose question, the two way analysis of variance did not yield any significant differences. The only $F$ greater than one was for the main effect of the rationale variable ($F = 2.27, \text{ ns}$), but again, the ordering of means for this variable was opposite to that predicted (mean reported time spent guessing about the experimenter's purpose was 2.39 for the rationale condition, and only 1.89 for the no rationale condition). The predictions that the absence of a rationale and the presence of a sensitization task cause subjects to spend more time guessing about the experiment's purpose were not confirmed.

In sum, the factors of rationale for experiments and sensitization were not demonstrated to have the predicted effect upon subject suspicion, apprehension, or time spent attempting to guess an experiment's purpose. In fact, the only finding that approached significance was that subjects tended to be more apprehensive without sensitization. This finding is difficult to interpret.

The second aspect of the analysis, correlating Questionnaire responses with degree of agreement and extremity of opinion on the Opinionnaire, was conducted in the following manner. Subjects were divided into two groups - suspicious and unsuspicious - based on their responses to the suspicion question on the questionnaire. Subjects who indicated that they were "slightly," "moderately," or "very" suspicious were placed in one category, and subjects who indicated they were not suspicious were placed in the other category. There were 18 suspicious and 18 unsuspicious subjects.
Suspicion level was first correlated with degree of agreement with the communication. Responses to the first four statements of the Opinionnaire were summed (item #3 being reversed) to obtain an "agreement" score for each subject. Mean agreement for unsuspicuous subjects was 18.00 and for suspicious subjects was 17.72 (the higher the score the more agreement). The point-biserial correlation for level of suspicion and degree of agreement was virtually nil ($r_{pb} = -.03$); the predicted negative correlation was not obtained.

The fifth statement on the Opinionnaire was analyzed separately as it allowed a partial check on the validity of the Questionnaire. It was expected that a subject might generalize the "ulterior motives" pronouncement made by the statement so as to be applicable to the experimenter. Since a suspicious subject by definition suspects an experimenter's motives, suspicion ought to correlate positively with agreement with this statement. However these two factors were not correlated ($r_{pb} = .09$). Evidence to support the validity of the Questionnaire as a measuring device was not obtained.

In order to test the level of suspicion-extremity of opinion hypothesis, Opinionnaire responses were rescored. A response to either of the endpoint intervals of the seven-interval Opinionnaire scale was given the highest score, and a response to the "neutral" position was given the lowest score:

\[
\begin{array}{cccccc}
\text{disagree} & 3 & 2 & 1 & 0 & 1 & 2 & 3 & \text{agree}
\end{array}
\]

Thus the higher the score the more extreme the opinion expressed (be it agreement or disagreement). Subjects' responses were summed across all five statements for the analysis. The predicted
negative correlation between suspicion level and extremity of opinion
was not, however, obtained ($r_{pb} = +.05$).

In summary, no performance differences were found between
suspicious and unsuspicious subjects.

DISCUSSION

This study was based on the principle that the stimulus to
subject suspicion and speculation is to be found in certain common
experimental practices rather than, for example, in a subject's level
of experimental and psychological sophistication. The effect that two
such practices - sensitization task and no rationale - have on arousing
suspicion and speculation could not be determined in this study.
Such results, although discouraging, do not necessarily invalidate the
principle, and research into the causes of suspicion and speculation,
because of the important implications, ought to continue.

The results of this study can perhaps be accounted for by the
lack of divergence between the values of the variables investigated.
First, the rationale-no rationale conditions may not have sufficiently
differed due to a procedural requirement. It was necessary to tell
subjects in the no rationale condition as part of the instructions for
filling out the Personality Impressions Scale whose personality they were
rating - otherwise they would have been unable to do the task. From
these instructions they may have deduced that the purpose of the
experiment had to do with personality impressions, and so in fact
have been left with the same information as subjects in the rationale
condition. Thus the two conditions may have had no difference in fact.
Secondly, the sensitization manipulation may not have been a particularly strong manipulation. The attitude survey-Opinionnaire relationship may have been made too subtle for subjects to comprehend, although this seems unlikely as inspection of the materials will reveal. In retrospect, it can only be said that it behooves an investigator when studying a new area to choose unsubtle, divergent values for his variables.

The honesty of subjects' Questionnaire reports is also, of course, debatable. The use of extensive post-experimental interviews and questionnaires as recommended by Levin (1961) and Orne (1962) probably are necessary to prevent responses which are too casual. Subject error may have obscured any treatment effects in the data as well as obviated the necessity of the correlational analyses.

The two studies reported in this paper, although predominantly unsuccessful in providing new knowledge in this area, will hopefully be instrumental in leading to new avenues of research. The study of demand awareness, in many ways meritorious, nonetheless may be restricted in generalizability. By studying all the hypotheses that subjects entertain rather than just the accurate ones, it may be possible to account for more instances of data contamination. It is of major importance, however, that future research in this area be directed towards determining the source of subjects' motivation to speculate and, if that source is found to be certain experimental
practises, towards enumeration of those practises which arouse subjects to speculate. If investigators adopt this course of action the artifactual problem can perhaps be solved. Speculation-arousing stimuli can be controlled or preferably eliminated from research designs entirely.

One strategy for determining the source of subjects' motivation to speculate might entail replication of some of the demand awareness studies with demand characteristics present and with demand characteristics absent. If removal of demand characteristics results not only in reduction of subject awareness but also in a reduction in subject suspicion and speculation, it can be concluded that motivation to speculate is not a function of extra-experimental factors such as knowledge of experimental deceptive practises, but a function of stimuli or events occurring within particular experiments, and specific to those experiments.

Some experimental stimuli which may be instrumental in arousing subject speculation and which therefore may be of interest for further research are as follows: 1) Stress. An experiment that is stressful for subjects may arouse them to speculate about its purpose. It is difficult to undergo stress without sufficient cause, and because a subject participating in a stressful experiment may experience misgivings about his having volunteered to participate, he may attempt to counteract these feelings by attributing a purpose to the experiment that would justify his participation. Stressful experiments therefore may increase subjects' motivation to have hypotheses. 2) The unexpected. Subjects may become aroused to speculate about an experiment in which they have been surprised,
especially when the surprise is an unpleasant one. For example, any large discrepancy between an experiment and information given subjects about that experiment at the time of recruitment may cause speculation, because of subjects' desire to anticipate future experimental surprises. Unannounced tasks may have a similar effect on subjects. Experimental practises which catch subjects off guard may put them on guard for the remainder of the experiment.

3) Inconsistency. Any inconsistencies in the totality of stimuli subjects receive about an experiment may cause them to generate hypotheses about that experiment. Tasks that appear to be unrelated to the experiment's stated purpose, contradictory utterances on the part of the experimenter, and procedures which conflict with subjects' knowledge of "accepted experimental practise," may induce subjects to speculate.

Although there are other possibilities, investigation of the speculation arousing effects of these three classes of variables would seem particularly worthwhile. Coupled with investigation of the source of subjects' desire to figure out an experiment's purpose, this type of research may lead, if not to solution of this artifactual problem, at least to its further clarification.
FOOTNOTES

1 Orne asserts that demand characteristics are present in all experiments, but his argument does not appear to support his claim. In making this assertion, he argues:

One of the basic characteristics of the human being is that he will ascribe purpose and meaning even in the absence of purpose and meaning. In an experiment where he knows some purpose exists, it is inconceivable for him not to form some hypothesis as to the purpose, based on some cues, no matter how meagre; this then will determine the demand characteristics which will be perceived by and operate for a particular subject. (1962, p.780)

These cues upon which "some hypothesis" is based are not necessarily the same cues which function to convey the experimenter's "demands" to the subject: according to Orne's formal definition demand characteristics cannot convey just any hypothesis, they can only convey the experimental hypothesis. In making this claim, Orne appears to have inadvertently extended his definition of demand characteristics to include any stimuli used by a subject as a basis for an hypothesis. Although it is certainly the case that all experiments contain stimuli which a subject might use as a basis for an hypothesis, these stimuli are not necessarily demand characteristics.

2 The hypotheses of subjects do not necessarily "randomize out." The demand awareness literature has shown that a considerable proportion of subjects are capable of coming up with the same idea about an experiment's purpose. A considerable proportion of subjects in an experiment free of demand characteristics may come up with the same "wrong" idea about its purpose.


4 Originally two communications were to be used in this experiment. The second communication argued that the United States would contribute very little that is new to the exploration of outer space. As U.S. astronauts were making their first landing on the moon at the time this experiment was being conducted, this communication condition was eliminated.

5 Independent judges were not used to classify subjects with respect to hypotheses. Page and Lumia (1968) report interjudge reliability to be very high when scoring for demand awareness, and their scoring procedure differs little from the one used in this study.
Although this subject's performance on the Personality Impressions Scale might have been affected by his hypothesis, it is unlikely, because he "didn't see how it could be an intelligence test." The performance of a subject whose attentions are divided may be affected, but this study is not addressed to that problem.

The apprehension question was included because of its possible relationship to suspicion and speculation.
BIBLIOGRAPHY


Orne, M. T. The nature of hypnosis: artifact and essence. Journal of Abnormal and Social Psychology, 1959, 58, 277-299. (a)

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


APPENDICES

A. Communication
B. Personality Impressions Scale
C. Opinionnaire
D. Opinionnaire - Experiment II
E. Questionnaire
F. Attitude Survey
APPENDIX A

Recently, there has been considerable discussion and debate about higher education. Here, we wish to focus on a single misapprehension that seems to have gained some acceptance by the public. This is the false conclusion that in the near future college graduates will need to undertake advanced graduate study before they are considered eligible for many jobs or positions open to them today. On the face of things, this contention may appear plausible; closer examination, however, reveals that it is incorrect, and that there is no reason to believe that the college degree is about to lose its status. Since beliefs about the status of college degrees are of considerable importance in determining educational policies, we should briefly review some evidence that shows that undergraduate college degrees will remain a sufficient qualification for employment.

To begin with the erroneous belief that advanced graduate degrees will be essential for employment in the future has developed as a result of misapplication of a familiar analogy: after all, was not the high school diploma sufficient some time ago, and is it not now necessary to have a college degree in addition to the diploma? In many ways, the diploma/college degree situation is indeed true. On the other hand, it provides no justification for drawing a similar analogy between the college degree and advanced graduate degrees. Many jobs today do require skills that are not taught in high school and are acquired in college. These same jobs, however, do not require graduate-level skills, and hence graduate work cannot become a realistic requirement for them.

A second, and equally important reason behind the insistence on college attendance that we find today has nothing whatsoever to do with education and skills, but is simply a result of the inability of the labor market to absorb too many people at any one time. College attendance increases the age at which a person enters the labor market, and this increase in minimum age is necessary to prevent the unemployment and low wages that would result from too many young people seeking employment at once. A very similar phenomenon is found in the trend toward an earlier retirement age. This too shrinks the size of the available labor force, and, again, reduces pressures that the economy would have been unable to bear.

Finally, let us note that additional pressures to shrink the labor force will appear in the future. These will be met by further reductions in the retirement age and not by raising the age of entry into the labor market. This second alternative is impossible because it would tend to eliminate the most productive, healthy, and recently trained segment of skilled workers and professionals.

Thus, it can be seen that the belief that college graduates will soon need advanced graduate training before they are considered eligible for many positions open to them today is completely unfounded. College attendance and graduation will continue to be meaningful and important employment criteria. Advanced degrees will continue to function as they do now: they will be required, as they are now, only of a few professionals engaging in highly specialized skills; as for example, physicians and top-level scientists.
APPENDIX B

PERSONALITY IMPRESSION SCALE

1. **Data on Rater**
   1. Age  ______
   2. Sex   ______
   3. Psychology course(s) in which presently enrolled (course # and section) __________________________
   4. Have you ever participated in a psychology experiment before? ________________________________
   5. Did you know anything about this experiment before you came? ____________________________
II. Instructions

On the next page there are a number of scales. Adjectives descriptive of personality traits appear at the endpoints of each scale. Each scale is marked into seven intervals. Place an "X" inside the interval that best reflects your impression of the author's personality.

To familiarize you with the use of the scale, here is an example. Suppose you were asked to give your impression of the author's height:

short _______ _______ _______ _______ _______ _______ _______ tall

If your impression is that the author is tall, place an "X" at the end of the scale closest to the word "tall", as follows:

short _______ _______ _______ _______ _______ _______ _______ tall

If your impression is that the author is short, place an "X" at the end of the scale closest to the word "short", as follows:

short _______ _______ _______ _______ _______ _______ _______ tall

If your impression is that the author is neither tall nor short, but of average height, place an "X" at the midpoint of the scale, as follows:

short _______ _______ _______ _______ _______ _______ _______ tall

In this fashion, place an "X" inside an interval in every one of the scales provided on the following page. Don't restrict yourself to marking only those intervals covered in the instructions — place an "X" in any one of the seven intervals you wish. Work fairly rapidly. Usually your first impression is the best guess. After you have finished marking your impressions, but not before, answer the question that appears on the last page of this form.
### III Scale

<table>
<thead>
<tr>
<th>Number</th>
<th>Description</th>
<th>Antonym</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Assertive</td>
<td>Submissive</td>
</tr>
<tr>
<td>2</td>
<td>Frivolous</td>
<td>Responsible</td>
</tr>
<tr>
<td>3</td>
<td>Cultured</td>
<td>Boorish</td>
</tr>
<tr>
<td>4</td>
<td>Stern</td>
<td>Kindly</td>
</tr>
<tr>
<td>5</td>
<td>Composed</td>
<td>Excitable</td>
</tr>
<tr>
<td>6</td>
<td>Eccentric</td>
<td>Conventional</td>
</tr>
<tr>
<td>7</td>
<td>Intelligent</td>
<td>Unintelligent</td>
</tr>
<tr>
<td>8</td>
<td>Adventurous</td>
<td>Cautious</td>
</tr>
<tr>
<td>9</td>
<td>Sociable</td>
<td>Reclusive</td>
</tr>
<tr>
<td>10</td>
<td>Practical</td>
<td>Imaginative</td>
</tr>
<tr>
<td>11</td>
<td>Mature</td>
<td>Immature</td>
</tr>
<tr>
<td>12</td>
<td>Rigid</td>
<td>Adaptable</td>
</tr>
</tbody>
</table>
IV. STRENGTH OF IMPRESSION

How strong was your impression of the personality of the author of this article?

quite weak ———————————————————— quite strong
OPINIONNAIRE INSTRUCTIONS

On the next page are statements, and each statement is followed by a seven interval scale. The left hand side of the scale is marked "disagree" and the right hand side marked "agree". The scale looks like this:

disagree ______________________ agree

Place an "X" inside the interval that best reflects how much you agree or disagree with the statement that precedes it.

To familiarize you with the use of the scale, here is an example:

Brunettes have more fun.

disagree ______________________ agree

If you are in strong agreement with the statement "Brunettes have more fun", place an "X" at the extreme right of the scale, as follows:

disagree ______________________ X agree

If you are in moderate disagreement with the statement, place an "X" in the next to last interval on the left side of the scale, as follows:

.disagree X ______________________ agree

If you are in mild agreement with the statement, place an "X" in the interval just to the right of center, as follows:

disagree ______________________ X agree

Mark each scale on the following page in the same fashion, depending on how much you agree or disagree with the statement that precedes it. Please mark all items.
OPINIONNAIRE

1. The article, on the whole, was a good one.
   disagree  
   agree

2. The evidence presented by the author was objective and factual.
   disagree  
   agree

3. The arguments used by the author were designed to play on the reader's emotions.
   disagree  
   agree

4. "The belief that advanced graduate study will soon be essential for employment is false".
   disagree  
   agree

5. The author of the article was sincere.
   disagree  
   agree
APPENDIX D

OPINIONNAIRE

1. The article, on the whole, was a good one.
   disagree [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] agree

2. The evidence presented by the author was objective and factual.
   disagree [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] agree

3. The arguments used by the author were designed to play on the reader's emotions.
   disagree [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] agree

4. The conclusions arrived at in the article are essentially correct.
   disagree [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] agree

5. The author obviously had ulterior motives for writing the article.
   disagree [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] agree
APPENDIX E

QUESTIONNAIRE

Please answer each of the following questions by placing an "X" in the appropriate interval. When necessary, explain your answer in the space provided below each question.

1. How careful were you in following instructions in this experiment?

| not at all | a little bit | moderately | very careful |

2. How suspicious were you of the experimenter's intentions?

| not at all | slightly | moderately | very suspicious |

3. How apprehensive were you about doing any of the tasks in this experiment?

| not at all | slightly | moderately | very apprehensive |

4. How much time did you spend attempting to guess the experiment's purpose?

| none at all | a small amount | a moderate amount | a lot amount |
ATTITUDES TOWARD HIGHER EDUCATION

Instructions: Each of the following statements is followed by a five interval scale ranging from "definitely false" to "definitely true". Place an "X" in the interval that most closely agrees with your opinion about the statement that precedes it. Please answer all items.

1. The increasing demand for higher education is a direct result of the inability of young people to find work.

   definitely | probably | uncertain | probably | definitely
   false     |     false |       true |     true

2. The purpose of education should be to teach people how to live rather than how to earn a living.

   definitely | probably | uncertain | probably | definitely
   false     |     false |       true |     true

3. Most students, had they any other choice, would not attend university.

   definitely | probably | uncertain | probably | definitely
   false     |     false |       true |     true

4. Universities will soon devote increased effort to students in their graduate programs.

   definitely | probably | uncertain | probably | definitely
   false     |     false |       true |     true

5. University education should be more concerned with the pressing issues of the day.

   definitely | probably | uncertain | probably | definitely
   false     |     false |       true |     true

6. A high school diploma is no longer sufficient to obtain a well-paying job.

   definitely | probably | uncertain | probably | definitely
   false     |     false |       true |     true