# DEMAND CHARACTERISTICS IN THE HINDSIGHT BIAS PARADIGM

by

Donald Sharpe

A thesis
presented to the University of Manitoba
in partial fulfillment of the
requirements for the degree of
Master of Arts
in
Psychology

Winnipeg, Manitoba

ISBN 0-315-37124-2

# DEMAND CHARACTERISTICS IN THE HINDSIGHT BIAS PARADIGM

BY

DONALD SHARPE

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF ARTS

I hereby declare that I am the sole author of this thesis.

I authorize the University of Manitoba to lend this thesis to other institutions or individuals for the purpose of scholarly research.

Donald Sharpe

I further authorize the University of Manitoba to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Donald Sharpe

ACKNOWLEDGEMENTS

The author would like to thank a number of individuals who provided assistance during the various phases of this thesis. First, and foremost, I wish to thank Dr. John Adair for his encouragement, guidance and painstaking efforts through all phases of this thesis. The committee members, Drs. James Forest and William Notz, are thanked for their helpful comments and their participation. Ms. Marcia Pollock is gratefully acknowledged for her tireless assistance with computer programing. Drs. Harvey Keselman and Cam-Loi Huynh provided advice on statistical matters. Dr. Jennifer Campbell made available the general knowledge statements employed within the study, Ms. Lorna Jakobson provided comments on the thesis proposal and Dr. David Koulack suggested precursors to the hindsight bias paradigm. Finally, I thank my friends and family for their support.

ABSTRACT

It has been suggested that when an event occurs, there is a tendency to see the event as having been predictable. Such a misperception may produce an exaggerated belief in a capacity to predict future events. The "knew-it-all-along" or hindsight bias has been investigated in the laboratory through a paradigm conceived to be functionally equivalent to the occurrence of the phenomenon in the real world. Stimulus material such as general knowledge statements are presented and subjects are asked to indicate their confidence in the validity of each. Subjects are then told whether each statement is true or not and are again asked to indicate how confident they are of the validity of each statement, ignoring the outcome feedback. Deviations from the original ratings given in the direction of the outcome feedback is seen as evidence of a hindsight bias.

A cognitive explanation for the bias argues that the outcome feedback alters the store of available knowledge, such that individuals are unable to assess or to ignore the influence of the outcome feedback. A motivational explanation proposes that dispositional factors (i.e. a need for positive self-presentation) mediate the bias. A demand

characteristic explanation implies that, for subjects in hindsight bias experiments, outcome feedback is one compelling cue to alter previous ratings. It is the latter explanation for the bias that the current study most directly addresses by manipulation of the cues that produce biased responding.

Subjects were asked to rate a number of general knowledge statements. This task was followed by a series of experimental manipulations of conditions employed in a previous hindsight bias study (Fischhoff, 1977). Fischhoff had claimed to have allegedly refuted the demand characteristic position by providing subjects with an explanation of the bias and forewarning to avoid the bias and finding no diminished bias for these instructions. In addition to variations in the completeness of Fischhoff's instructions, one form of the instructions attempted to induce an expectation of a reversed hindsight bias. After again rating a number of statements, additional instructions were given to subjects requesting that they intentionally rate subsequent sets of statements in either a biased or a non-biased manner. Motivational factors, assessed between the administration of the first and second sets of general knowledge statements, were correlated with subjects ratings of the statements. Responses on a post-experimental questionnaire were evaluated to assess the extent to which biased responding was conscious.

Providing subjects with various forms of explanations and forewarnings of the bias had no effect. The bias was altered, though, by demands for intentional biased or non-biased ratings, an effect that interacted with the explanation given for the bias. Specifically, subjects could produce or not produce the bias on demand, with the direction of the bias consistent with the form of the explanation that had been given. Both samples that had received or had not received an explanation for the bias recognized the role of the outcome feedback in influencing their ratings. For the latter, biased responding was related to awareness as to the purpose for having been provided with outcome feedback. Some limited support for the involvement of motivational factors was also found. All of these results were discussed in terms of explanations for the bias and consequences for research within the experimental paradigm.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE PAGE

# INTRODUCTION

How individuals select one outcome from a number of alternatives has been of considerable interest to psychologists. The one consistent finding from this diverse literature is that the decision making process is far from perfect and is susceptible to systematic misinterpretations. One such systematic error in relation to the appraisal of the decision making capacity has been termed the hindsight bias. An individual is presented with a problem and a solution is selected. In time, the correctness of the solution selected becomes apparent. If, with this knowledge, the individual reports an exaggerated confidence in an a priori recognition of the correctness of the solution selected, then the individual is said to have demonstrated a "knew-it-all-along" or hindsight bias.

While anecdotal reports suggest that a hindsight bias may commonly operate, study of the phenomenon and the antecedent variables affecting it requires a laboratory paradigm that recreates the bias under controlled conditions. The paradigm developed and assumed to meet these objectives is initiated by providing respondents with a number of problems to consider. One or more possible solutions are offered for

each problem, and the subject is to indicate which of the solutions is correct and the degree of their confidence in this judgement. For instance, one might be asked to indicate the extent of confidence in a general knowledge statement as being factually correct. After a period of time has elapsed, the validity of the statement is revealed. Subjects are implored to ignore this outcome feedback while rerating the statement for correctness. Those subjects who tend to give higher ratings for correct solutions and lower ratings for incorrect solutions are said to have been influenced by the feedback provided during the second examination of the items and to have shown a hindsight bias. In addition to general knowledge statements (Campbell & Tesser, 1983; Fischhoff, 1977; Wood, 1978), historical events (Fischhoff & Beyth, 1975), pregnancy testing (Pennington, Rutter, McKenna & Morley, 1980), word pairs (Campbell & Tesser, 1982) and medical diagnoses (Arkes, Sarville, Wortmann & Harkness, 1981) have also been employed to show a hindsight bias.

An alternative means of demonstrating the bias, without forcing subjects to rerate the same set of items twice, is to have subjects rate the items for the first time after having received outcome feedback. For instance, the first presentation of a statement to a particular respondent, as described above, could be with the outcome feedback

provided. The rating given by this sophisticated subject for the statement will then be compared with the mean rating given for the statement by respondents who rated the statement without the benefit of outcome feedback. Presentation of a set of statements only one time to a respondent, with outcome feedback, is termed the hypothetical condition. Presentation of a set of statements twice to a respondent, the second time with outcome feedback, is termed the memory condition. While a more substantial bias has been found within the hypothetical condition (Campbell & Tesser, 1983), this difference is only a matter of degree.

Attempts to demonstrate the hindsight bias, also in the laboratory, with historical events and with problem solving tasks have produced results similar to those found with general knowledge statements. Fischhoff and Beyth (1975) presented subjects with a number of possible outcomes to then President Nixon's visits to China and Russia. The amount of time between the ratings of the possible outcomes of the visits, and descriptions of the visits, were varied. Subjects were also questioned during the second set of ratings as to whether or not they believed that the outcomes specified had indeed occured. Subjects asked to reproduce the same probabilities that they had previously given composed the memory condition. Subjects absent for the

first testing were asked to reconstruct the ratings that they would have given prior to the event if they had been queried at that time, thereby creating a control sample. Both groups of subjects produced results consistent with the hindsight bias phenomenon, although somewhat less so for events that had not occurred.

Fischhoff (1975) reported three experiments in which an historical event was described and possible outcomes were presented. 'After' subjects were told that one of the outcomes was true. 'Before' subjects were not provided with such information. It was found that 'After' subjects rated the outcomes which they had been told were correct as more likely to have occurred, compared to ratings given by 'Before' subjects for the same events. In a second experiment, subjects were instructed to answer as if they had not been provided with the correct outcomes. Similar results to the first experiment were produced, leading Fischhoff to conclude that "subjects are either unaware of outcome knowledge having an effect on their perceptions or, if aware, they are unable to ignore or rescind that effect" (p.295). In a third experiment, subjects were asked to respond as a peer might. Results were similar to those found in the first and second experiments. It was concluded that subjects were unaware of the effect of outcome knowledge on judgements because varying the instructions given did not

alter performance. This conclusions was made although no direct assessment of the extent to which subjects were aware or were conscious of the effects of outcome knowledge was made.

Goranson (1985) explored the hindsight bias phenomenon with a variety of problem solving tasks. Subjects were provided with the solutions to problems and then attempted to estimate how difficult it would have been to solve the problems if the solutions had not been given. Ratings were on the basis of estimates of the amount of time that would have been required or the number of errors that would have been made. These values were compared with the actual performance of subjects who attempted the problems without receiving the correct solutions. In a number of the tasks, informed subjects actually attempted the problems and their performance was compared with the performance of naive subjects. Not surprisingly, ratings of difficulty were lower, and actual performance was better, for subjects provided with the correct solutions. Goranson argues that the finding that informed subjects judged the difficulty of the problems as less, or completed the problems more rapidly than naive subjects, extends the generalizeability of the hindsight bias effect. How long an informed subject should have waited until a period of time had passed that would be equivalent to having solved the problems without outcome

knowledge is, from the perspective of the subject, unknowable.

## Hindsight Bias in Non-Laboratory Contexts

Concerns over external validity have lead to attempts to demonstrate the bias in areas of human judgement distinct from the laboratory. Pennington, et al., (1980) queried women who suspected that they might be pregnant. Half of the women gave a rating on how likely it was that they were pregnant five days prior to, and then immediately after having received the results from a pregnancy test. A second group of women gave their first ratings immediately after finding out the outcome of their pregnancy test. These women were also asked to reconstruct how likely they had thought that they were indeed pregnant five days prior to the pregnancy test. The hindsight bias was found only with the second group of women. Pennington et al., concluded that the experiment had provided only partial support for the hindsight bias phenomenon. In addition, women who had received a positive test result, not surprisingly, gave higher estimates than women whose test result proved negative. These women had had their suspicions confirmed, and were therefore a subsample who were "biased to seeing this outcome as more likely in hindsight" (Pennington et al., 1980, p.323).

Arkes, et al. (1981) presented physicians with a number of symptoms and then asked them to express their confidence in each of four possible diagnoses as being correct. Hindsight subjects were told that the patient had previously been diagnosed as having one of the four illnesses presented. Foresight subjects received no information on previous diagnoses of the patient. Only one of the four diagnoses resulted in a hindsight bias, although slightly more than half of the hindsight subjects gave higher probability estimates to the disorder for which they had been led to believe the patient had been previously diagnosed.

Pennington (1981) examined the hindsight bias in relation to a firemen strike. Time of assessment and whether subjects were restricted to possible future outcomes or could generate their own outcomes to the strike were also varied. Only two of the five specified outcomes produced a hindsight bias effect. When subjects were asked to generate their own outcomes, little evidence for the bias was found. It was suggested that the recall of the foresight state was enhanced through this procedure.

Whereas there is some evidence for the hindsight bias outside the laboratory, these demonstrations have not produced as consistent an effect or as robust a phenomenon as found in the laboratory with general knowledge statements

and historical events. This is of concern since the paradigm was developed to be analogous to the bias as observed outside the laboratory.

## Characteristics of the Stimulus Material

The nature of the items to be rated has been predicted to influence the appearance and the size of the hindsight bias. Fischhoff and Beyth (1975) found outcome feedback that labelled a statement as true produced a more powerful bias than outcome feedback that labelled a statement as false. Fischhoff (1977) found that implausible statements labelled as true and plausible statements labelled as false were most likely to produce a strong bias. In contrast, the second study of Wood (1978) compared statements from his first study that were most strongly correlated with the bias with those statements that were correlated with the bias the least. Regardless of which set of statements were responded to, the hindsight bias was strongly replicated. This was interpreted as suggesting that, while the size of the effect may vary according to whether the outcome feedback labels the statements as being true or false, other characteristics of the statements do not significantly influence the size of the effect.

On the assumption that highly meaningful word pairs are capable of generating numerous associations, Campbell and

Tesser (1982) compared word pairs of high and low meaningfulness. Subjects were twice presented with a list of word pairs. For the second presentation of the word pairs, the first word of a pair was presented with either its match from the list or a different word of the same level of meaningfulness, and with or without feedback as to the status of the second word. It was found that there was a strong relationship between occurrence of a hindsight bias and whether a word pair was highly meaningful. Although confounded with item meaningfulness, an index of item ease was also calculated. This was defined as a deviation in the direction of the correct answer on the first set of ratings from the center portion of the response scale, the portion which indicates no knowledge of the correctness of the item. The item ease index and the magnitude of the bias were strongly negatively correlated, suggesting that to the extent that "persons are certain of the correct response to an item without the benefit of feedback, it is clearly more difficult to demonstrate the bias" (p.19). Wood (1978) stated that "individuals who know very little [specifically to the item] prior to feedback have a better chance of demonstrating the effect, since ignorance is a precondition for obtaining the effect" (p.352). Whereas the evidence suggests that the hindsight bias effect may not be item specific, it may indeed be influenced by both the content of the feedback provided and the difficulty of the items employed.

## Procedures to Reduce the Bias

A number of attempts have been made to systematically reduce or eliminate the hindsight bias. Cognitive theorists argue that only through a restructuring of the available information store can the hindsight bias be systematically eradicated. Slovic and Fischhoff (1977) significantly reduced the bias by presenting subjects with the results of scientific experiments and then asking subjects to consider "'Had the study worked out the other way, how would you explain it?'" (p.548). Pennington (1981) found that providing less detailed and less informative outcomes reduced the bias. Similarly, whether the consequences of an outcome were perceived as being positive or negative influenced the occurrence of the bias in relation to ratings of nursing behavior (Mitchell & Kalb, 1981). One strategy employed has been to distance the subjects from the events being rated. For instance, Fischhoff and Beyth (1975) found that for events perceived not to have occurred, the bias could be reduced through increasing the time period between the first and the second ratings of the outcomes.

In a number of instances in which it was predicted the bias would be reduced or eliminated through some manipulation, such was not found. Wood (1978) delayed re-rating of general knowledge statements previously given until after presentation of all of the outcome feedback.

The results from such a procedure did not differ from those found when the feedback was presented at the time of re-rating. Fischhoff (1982) cites a number of studies which indicated that expertise in the content area of the items to be rated does not reduce the bias. For instance, Wood (1978) did not find that subjects with the greatest ability in relation to identifying general knowledge statements produced any less bias than those individuals with lesser ability.

In the first of two experiments, Hasher, Attig and Abla (1979) had subjects examine statements divided into two lists of ten. One list was labelled as containing true items, the other false items. After one group of subjects had studied the two lists, the experimenter then apologized and indicated that those statements labelled as false were actually true, and that those labelled as true were actually false. Another group received only the feedback without the disconfirmation manipulation. A third group received no feedback. All of the subjects were then asked to rerate the statements exactly as they had done prior to receiving outcome feedback. The disconfirmation procedure was not found to diminish the influence of the hindsight bias. Statements were rated in a biased manner according to the outcome feedback received, whether that feedback was the original or the corrected outcome feedback. It was argued

that the manipulation had not been sufficient to suspend belief in the credibility of the feedback, as subjects considered having first been provided with incorrect feedback as an honest mistake.

In a second experiment, subjects in a 'Disconfirmed/Wrong' condition were told that the feedback sheets were incorrect, and that they were to ignore the feedback and to rerate the statements as before. No explanation was given for why incorrect feedback had been given. Subjects in a 'Disconfirmed/Mistake' condition were offered justification for having been provided with incorrect feedback, that it was a mistake and was not important for the purposes of the study. Subjects in the latter condition produced a somewhat diminished hindsight bias, replicating the first experiment, but subjects in the former condition did not demonstrate a hindsight bias.

Fischhoff (1982) views the attempts of Hasher et al., (1979) and Slovic and Fischhoff (1977) as the only reliable means of eliminating the hindsight bias, presumably through a cognitive restructuring of the store of available information about the item. This perspective demands that the feedback provided be discredited, either through reducing confidence in the source of the outcome feedback or through instructions to consider alternatives to the outcome feedback presented. According to this view, mere

exhortations to avoid the bias will not succeed. Furthermore, such a position may explain the diminished bias found by Pennington (1981) and by Mitchell and Kalb (1981). In both studies, it was the information available to respondents that was varied. Aside from the theoretical consequences, such a perspective has implications for how one might ameliorate defective hindsight judgement in practical settings. For instance, Janoff-Bulman, Timko and Carli (1985) suggest that victims of violent crime are often blamed, in hindsight, for their predicament.

## EXPLANATIONS FOR THE BIAS

Several explanations have been offered for the hindsight bias. Specifically, cognitive, motivational and demand characteristic explanations have been proposed, and evidence for each has been gathered.

### Cognitive explanation.

The explanation for the hindsight bias phenomenon which has gained the most acceptance is the cognitive or the processing of information explanation. When the subject receives outcome feedback, this knowledge is assimilated into the previous store of information concerning the item. "The retrospective judge attempts to make sense, or a coherent whole, out of all that he knows about the event" (Fischhoff, 1975, p.297). The resulting change in judgements of confidence are viewed as a product of this assimilation. Similarly, from an attributional perspective, the "apparent inevitability that events seem to accrue when viewed with hindsight may result...from the explanatory framework the individual has generated in reflecting upon that event." (Ross, Lepper, Strack and Steinmetz, 1977, p.826).

Fischhoff and Beyth (1975) believe that this assimilation or learning from experience must be partially conscious. In contrast, Campbell and Tesser (1982), in summarizing the literature, suggest that subjects are "largely unaware that receipt of outcome knowledge has altered their perceptions" (p.22). Wood (1978) contends that the failure to find a difference in the size of the effect with instructions to rate items as one's peers might, suggests that assimilation is automatic and does not involve conscious mediating factors. While assimilation is presumed to be automatic, the manipulations that have been employed to reduce the effect have demanded thoughtful activity on the part of respondents. This activity may serve to disrupt the product of these automatic processes rather than altering the effect itself (Campbell & Tesser, 1982).

Attempts to eradicate a hindsight bias involve restructuring of the store of available knowledge. Since restructuring must occur after the automatic retrieval of cognitions has commenced, restructuring must function to reverse the assimilation process. Campbell and Tesser (1982) contend that "It is inferentially inappropriate to argue that because persons required to list contradictory reasons show a reduced effect that selective retrieval of reasons is implicated in the production of the effect" (p.8). Further, individuals who are required to list alternative reasons for

an outcome are operating under an experimental demand to employ caution, a demand that is not present within comparison groups (Campbell & Tesser, 1982).

After the outcome information has been assimilated into the previous store of knowledge, the cognitive heuristics utilized to make subjective probability estimates are then employed. For instance, if prior information supports the outcome feedback given, supportive reasons may be more easily accessed than contradictory reasons. Alternatively, individuals may make estimates of correctness from an initial starting value that is adjusted through the assimilation of outcome feedback to produce the final solution. This final value may be biased as a function of both the initial starting value and faults in the adjustment process. Regardless of which cognitive heuristic is employed, the cognitive explanation demands that there be some form of biased retrieval of prior knowledge (Campbell & Tesser, 1982), this biased retrieval producing the 'knew-it-all-along' effect.

Motivational explanation.

Motivational factors have been proposed to mediate the occurrence of a hindsight bias. One motive that may be involved is the need for control or predictability. The desire to predict events in the environment is tied to the

human need for control. Believing "that one knew all along which outcome was correct implies competency in one's interactions with the environment" (Campbell & Tesser, 1983, p.607). A related motive that might be involved is that of self-presentation. It is suggested that "to impress others or to save face" (p.607), it might be of benefit to believe that knowledge of the correct outcome was possessed without needing to receive feedback.

A number of theoretical arguments have been offered to refute the involvement of motivational factors within the hindsight bias. Fischhoff (1977) argues that social desirability cannot be involved in the production of the hindsight bias since the description of the task he employed was that of a test of memory, and was therefore not evaluative. Wood (1978) suggests that explaining the purpose of the study as a test of general knowledge should reduce attempts to "consciously please the experimenter" (p.347). The argument that assessment of one's memory or one's general knowledge store is not sufficient to induce evaluation apprehension is dubious.

The failure to find a difference between peer and self-ratings in the Wood (1978) study is viewed by Fischhoff (1982) as strong evidence against a self-presentation explanation for the effect. It is hypothesized that the self-rating instructions should have aroused greater

motivational factors than instructions to rate the items as one's peers might. Since the respondent knows that it is his knowledge that is being tested, and not the knowledge of his peers, one has to question the assumption that peer ratings are different from self-ratings in relation to self-presentation concerns. Moreover, Campbell and Tesser (1982) contend that both Fischhoff (1975) and Wood (1978) data suggests a greater effect with self-ratings than with peer ratings, which would support a self-presentation explanation for the bias. However, no statistical analyses were performed on the Fischhoff data, and a significant difference was obtained only in the second of the two studies of Wood.

Campbell and Tesser (1982) employed an importance manipulation such that to the "extent subjects believe it is important to know (or to appear to know) the correct answer, both effectance and esteem maintenance motives should be aroused" (p.12). Subjects in the high importance condition were told that the experimental task, learning pairs of words, was related to a number of desirable traits such as intelligence, skill in social situations, and academic performance. Subjects in the low importance condition were told that past research indicated that their performance on the task would not be related to those constructs.

The importance manipulation was unsuccessful. A post-experimental questionnaire indicated that subjects assigned to the high importance condition were not differentiated from subjects in the low importance condition on items such as whether the skill being examined was related to constructs of some consequence. It was found that subjects in the high importance condition did put more effort into the experimental task, and that for highly meaningful word pairs, subjects in the high importance condition did demonstrate more of a hindsight bias. A significant correlation was reported such that subjects who indicated on the post-experimental questionnaire that the task was of greater importance to them were also more likely to produce a hindsight bias. Campbell and Tesser suggest that their manipulations of the motivational factor was ineffective primarily because of the time period between the administration of the experimental manipulation and the post-experimental questionnaire. Because of this difficulty, the conclusiveness of their findings are limited.

In another attempt to support the motivational perspective, Campbell & Tesser (1983) examined the relationship of self-presentation and desire for predictability motives to the hindsight bias. Subjects rated their confidence in the correctness of one of two sets of general knowledge statements. After completing

self-presentation and desire for predictability measures, subjects rated both sets of items with outcome feedback provided, thereby producing a hypothetical and a memory condition. It was found that the motivational variables were related to the production of the bias. This relationship was found more so with memory instructions than with hypothetical instructions, though it was not a statistically significant difference.

It must be noted that Campbell and Tesser (1983) chose to have their subjects complete the personality measures in the time period between the first and second presentations of the general knowledge items. Completion of the measures at such a time might have served to induce motivational factors. In addition, Campbell and Tesser cautioned that their research design was correlational, such that the possibility of a "third, unmeasured, causal factor... [which] caused some subjects to score high on the two motives and to exhibit a large bias" (p.617) cannot be ruled out.

The Campbell and Tesser studies reflect the difficulties in demonstrating a motivational antecedent to the hindsight bias effect. Except from a strictly cognitive perspective, it is intuitively difficult to rationalize motivational factors as not implicated in the production of the bias. Yet, the laboratory manipulations of motivational factors

have generally not been effective. Experimental attempts to induce motivational factors, such as a desire for predictability and a need for positive self-presentation, have been transient in form. More correctly, these factors should be conceived as relatively stable and enduring traits. Nevertheless, the correlational findings presented do suggest that the effect may not be strictly cognitive in origin or in form.

Demand characteristic explanation.

A third plausible explanation of the hindsight bias effect may be the demand characteristics present within the experimental paradigm employed to demonstrate the bias. Demand characteristics are the cues in the experiment that convey a hypothesis to the subjects. Orne (1970) argues that behavior in any experiment is a function of both experimentally manipulated variables and demand characteristics. It is from these latter cues that subjects form some concept of the purpose of the study. Once acquired, this concept guides subjects' behavior. Demand characteristics need not be obvious to be effective, yet they are capable of producing behaviors which, although artifactual, may appear to be experimentally derived (Orne, 1970).

From a phenomenological perspective, the hindsight bias paradigm provides both the incentive and the opportunity for subjects to respond on the basis of demand cues. In particular, several events occur within the paradigm that demand explanation. Although explicitly instructed to ignore the outcome feedback, subjects recognize that this feedback has been provided for some purpose. In almost any conceivable situation, it is indeed inappropriate not to use feedback to improve performance. Since subjects are required to re-rate the same set or a similar set of statements twice, the second time with outcome feedback, it must be apparent that some change in the second set of ratings is expected. Undoubtedly, it is the outcome feedback that indicates to subjects the direction that the change from the first to the second set of ratings should take. Furthermore, the difficulty of the statements to be rated suggests that it is not the knowledge possessed by subjects that is of interest to the researcher. This also serves to highlight the change in the two sets of ratings as the behavior of interest.

The experimenter intends for the outcome feedback to unconsciously influence subjects' ratings. Subjects, on the other hand, may well consciously employ the outcome feedback as a cue as to how to produce the behavior that is perceived to be desired by the experimenter and acceptable within the

context of the experiment. Prior to conceding the intentions underlying subjects' behavior, one must know that the subjects' perceptions and the experimenter's intentions are the same (Adair & Schacter, 1972). It is indeed the case that individuals are often required to make assessments of the accuracy of propositions encountered in the real world, that feedback on the accuracy of their assessments will become apparent, and that this assessment may then be distorted as a function of this feedback. The difference between the real-world situation and the experimental analogy is that the participants in the latter are aware that they are involved in a psychological experiment, that the researcher is interested in the production of a particular behavior as a function of the manipulations in the study, and that the most apparent manipulation is that feedback information is being provided. Subjects need only then decide if they are indeed going to be 'influenced' by the manipulation and to produce responses indistinguishable from those that are the product of a true psychological effect.

Although a demand characteristic explanation for the bias has apparently been dismissed by most researchers, this rejection is based largely on argument or on weak empirical evidence that can just as readily be interpreted in favor of a demand characteristic interpretation. Campbell and Tesser

(1982) argue that the instructions create experimental demands for accurate recall or for ignoring outcome information rather than for deviating from ratings previously given. Unless subjects' perceptions of these instructions are examined, the precise experimental demands being induced by these instructions is unknowable.

That subjects deny the influence of the experimental manipulations is not a defense against the argument that the instructions are being used to assist them to know how to respond. Gilovich (1981), examining biased information processing, noted that, following debriefing interviews, "Subjects unanimously denied that the manipulations had had any influence on their decisions. Subjects tended to maintain that they noticed the various manipulations and thought that some people might be influenced by them, but that they themselves were not" (p.807). Orne (1962) has noted that there are powerful demands on subjects not to reveal that the purpose of the experiment has been 'guessed', thereby invalidating experimental participation.

Finally, the robustness of the phenomenon across instructional sets and stimulus materials is viewed as conflicting with a demand characteristic interpretation (Wood, 1978). There are other robust phenomena, such as the classical conditioning of attitudes, for example, that are the product of demand awareness (Page, 1973).

The alleged empirical evidence against a demand characteristic explanation for the bias is also not compelling. Wood (1978) concluded that the occurrence of a hindsight bias within the peer condition in his first study, and in the debiasing condition in the second study of Fischhoff (1977), ruled out demand characteristics as a plausible source of the bias. Wood presumed, without providing supporting evidence, that respondents instructed to rate items as others might, are operating under fewer experimental demands. It is debatable whether there are fewer experimental demands when subjects supposedly immerse themselves in the role of another. Subjects are well aware that it is their behavior, rather than the behavior of their peers, that is of interest.

The Fischhoff (1977) study was conceived as a direct test of the role of demand characteristics in the hindsight bias paradigm. Subjects were placed in either a hypothetical condition, a hypothetical warning condition, in which subjects were instructed to devote attention to the task, or a hypothetical debiasing condition, in which the subjects were instructed in the nature of the bias and were asked to consider alternative outcomes. Fischhoff's debiasing instructions are reproduced in Appendix A. Such a procedure is similar to the process debriefing technique successfully utilized by Ross, Lepper and Hubbard (1975) to ameliorate

faulty self and social perceptions. Nevertheless, the hindsight bias effect was found for all three conditions, for both true and false labelled alternatives.

However, instructions to the hypothetical debiasing subjects concluded with the disclaimer to "'be careful not to overcorrect and sell yourself short by underestimating how much you would have known without the answer'" (p.335). This being the final instructions given to subjects might have placed a powerful demand in the minds of subjects as to what the experimenter desired their behavior to be. At best, such instructions only may have served to neutralize the previous description of the bias and the instructions to avoid it.

At worst, Fischhoff's instructions might have heightened demands to utilize the feedback information to alter one's confidence ratings. Some support for this view is found in the observation of a small hindsight bias (though, not significantly) demonstrated by subjects in the hypothetical debiasing condition. Nonetheless, the finding of a hindsight bias for all three conditions in Fischhoff (1977) does not dismiss demand characteristics as a source of the effect.

## STATEMENT OF THE PROBLEM

The demand characteristic, cognitive, and motivational explanations of the hindsight bias are consistent in the conclusion that outcome information influences the ratings of confidence. It is the form of the influence that causes subjects to alter their ratings that is contested. From the cognitive perspective, subjects may or may not recognize that the outcome feedback has been provided to influence their re-rating of statements, but even if they do, the process operating within the hindsight bias paradigm is not a function of this recognition. Rather, the effect within the experimental paradigm is seen as analogous to the bias as anecdotally reported, an effect that is the product of faulty human judgement.

From the motivational perspective, the experimentally generated bias is the product of a psychological phenomenon mediated by personality variables. Individuals with a high need for social approval, a high need to maintain self-esteem or a high need to control self-presentation will attempt to produce behavior that would appear to suggest that the correctness of the statement was known prior to receiving feedback. These personality variables are not

related directly to participation in the study, except to the extent that participation is a social event, inducing factors present in any social encounter.

The demand characteristic perspective argues against the laboratory paradigm as producing a true hindsight bias. Instead, an artifact is formed which resembles the hindsight bias in that it is the result of a cognitive appraisal of the outcome feedback and mediated by motivational factors to produce behavior compatible with the subjects' interpretation of the experiment. Instructions to consider alternative outcomes are conceived of as reducing the bias by reducing the experimental demands to use outcome information as the only appropriate means of determining how one should respond.

It is only when one considers the process that produces the bias, whether the bias is a product of a psychological phenomenon or of experimental participation, are the three perspectives distinguishable. Attempts to understand the source of the bias from merely examining the responses generated is not of value. This is because the demand characteristic, cognitive and motivational explanations of the bias do not disagree on the product of the bias, that ratings of confidence increase with outcome feedback. Disagreement centers on the antecedents of the effect. What is needed then is a manipulation of these antecedent

conditions, and then an examination of the effects of the manipulations on ratings given in relation to the three explanations offered for the bias. A more appropriate manipulation would eliminate certain problems of the type found in Fischhoff's (1977) examination of demand characteristics. A phenomenological examination of how respondents view the experimental situation would determine the extent to which subjects are aware of the demands in the paradigm and the extent to which production of the bias is conscious. An assessment of individual difference factors, as related to a motivational source of the bias, needs also to be made.

Asch (1952) had similar difficulties with a paradigm that resembles the hindsight bias paradigm. Influence of majority opinion, rather than the hindsight bias, was the issue of interest. The paradigm originated with Moore (1921) who had presented subjects with pairs of statements related to judgements of grammatical correctness, ethical infringements, and consonance of musical chords. For instance, two statements regarding two different, improper behaviors were presented. Subjects rated which behavior was ethically less appropriate. The same statements were rated again a few days and a few months later, the second rating with information as to how the majority had rated the statements. It was found that the judgements of the

statements were modified from the first set of ratings to the last set according to the direction of the majority opinion. Asch (1952) contested that researchers using the paradigm had "presupposed the process and considered the sole remaining problem to be that of measuring it" (p. 404). It was suggested that subjects may have perceived the statements as "trivial" and as "lacking in reality" (p. 405). Subjects would then appear to be swayed by the majority opinion, although this behavior would then "not have much in common with the problem with which the investigator started" (p. 405). Asch concluded that one explanation for the effect might be what was later termed the demand characteristics in the paradigm.

> Once the subject has accepted the task he feels the need to arrive at a judgement. Not having a clear basis to go on he leans on the clues the experimenter has placed in his path. But his concern may no longer be that of reaching a clear conclusion but to respond in a way to escape censure or ridicule. The result may be that his expressions of judgements do not carry conviction to himself and no longer represent actual evaluations. (p. 407)

Furthermore, Asch suggests that while researchers took care "to hide the purpose of the investigation from their

subjects", he bemoaned that "researchers...have not inquired into the phenomenol aspect of the situation" (p.410).

The present study addresses the problem of assessing the source of the bias through manipulating the instructions provided to respondents. This was done to determine the extent to which the feedback, in particular, and the experimental situation, in general, are the sources of the bias. The primary objective was to demonstrate the involvement of demand characteristics within the hindsight bias paradigm. The results from the Fischhoff (1977) debiasing condition is the most often cited justification for dismissal of the demand characteristic interpretation of the effect. This may well be an unwarranted assumption.

The instructions provided by Fischhoff to respondents were manipulated to provide a fair test of the involvement of demand characteristics. This served to assess whether the experimental paradigm is an acceptable means of examining the bias or whether results obtained are artifacts of experimental participation. Two manipulations of these instructions took place. The first was to make the debiasing instructions more precise. The latter portion of the Fischhoff instructions, which could convey an alternative meaning to subjects, were deleted so as to provide a more explicit explanation for the effect. The 'complete' instructions provided by Fischhoff to subjects were replaced

by a 'shortened' version of instructions. The second manipulation was that the descriptors of the bias which suggest that ratings move in the direction of the outcome feedback were replaced by descriptors that imply the bias is the product of ratings that move away from the direction implied by the outcome feedback. The 'knew-it-all-along' effect was thus replaced by what might be called the 'I-never-knew-that' effect. Except for variations in the length of the explanation and in a small number of descriptors that relate to the direction ratings take under the influence of the bias, the instructions to subjects were comparable to those provided by Fischhoff (1977).

After the debiasing instructions were given, and a set of statements with outcome feedback were rated, an additional set of statements were presented for rating. Preceding this final set of statements were instructions for each statement to be rated either under or not under the influence of the bias. In this way, experimental demands for subjects to rate statements in a biased or a non-biased way were made more explicit. From the cognitive or motivational perspective of the bias as a psychological phenomenon, one would not expect a difference between statements rated under either biased or non-biased rating instructions.

The second objective was to phenomenologically determine the extent to which subjects were conscious of biased

responding. Post-experimental questionnaires have been employed in the hindsight bias literature but only as manipulation checks. The post-experimental questionnaire employed in this study functioned to assess the awareness of the subjects to the influence of the feedback information. Such data is not only useful in light of the cognitive perspective that the bias is unconscious, but assists in determining the extent to which demand characteristics may be a plausible explanation for the phenomenon. The post-experimental questionnaire responses of two groups of subjects were examined. The first group were 'no instruction' subjects, subjects who did receive debiasing instructions. Since these subjects had not experienced any debiasing instructions or any other manipulations, a measure of awareness within a group of subjects who were not forewarned of the bias was provided. The second group were a subsample of subjects from the factorial portion of the study. This subsample was selected to be the same size as the 'no instruction' sample and so to be balanced across the experimental manipulations. Since subjects from this subsample had experienced a form of the debiasing instructions, an assessment of the impact of these manipulations on experimental awareness was made possible by comparing the post-experimental responses of the first and second groups.

The third objective was to examine the involvement of motivational factors. While a motivational perspective must necessarily accommodate the involvement of cognitive factors, the strict cognitive position denies motivational factors as both a source of, and as an influence on, the phenomenon. A number of personality measures were administered, and the relationships between these indicies and the measures of the hindsight bias were assessed. These scales were the Self-Monitoring Scale (Briggs, Cheek, & Buss, 1980), the Texas Social Behavior Inventory (Helmreich & Stapp, 1974), and the Marlowe-Crowne Social Desirability Scale (Reynolds, 1982). The three scales were compared with demonstration of the hindsight bias. The extent to which an individual is capable of detecting cues to produce appropriate social behavior (self-monitoring), has a high or a low perception of self-worth (self-esteem), or is in need of social approval (social desirability) might intuitively appear to be related to a hindsight bias that is motivationally based.

The following specific hypotheses were posed:

(1) Subjects assigned to the 'knew-it-all-along' condition should demonstrate the hindsight bias, whereas subjects in the 'never-knew-that' condition should demonstrate a reversed hindsight bias. To the extent that biased responding is the product of demand characteristics,

subjects should respond according to the explanation of the bias that is provided. Thus, a 'reverse' bias, is predicted for 'never-knew-that' subjects. Subjects in the 'knew-it-all-along' condition should demonstrate a hindsight bias in the direction of the outcome feedback, replicating the findings of Fischhoff (1977).

(1a) Because it was expected that the two concluding paragraphs of Fischhoff's debiasing instructions were contradictory to prior debiasing instructions to avoid the bias, less of a bias was predicted for subjects in the 'shortened' instruction condition than for subjects who received the 'complete' instructions.

(2) Instructions to rate the last set of statements, in a biased or in a non-biased way, was predicted to result in ratings consistent with these instructions (i.e., biased/not biased) and with the description of the bias given previously ('knew-it-all-along'/'never-knew-that'). To the extent that the bias can be influenced by instructions provided to subjects to rate statements in a biased or in a non-biased way, the greater the support for the position that the laboratory demonstration of the bias is the product of an artifact, specifically experimental demands to respond in a biased manner.

(3) The 'no instruction' subjects, whose responses on the post-experimental questionnaire indicate awareness as to the reasons for having been provided with outcome feedback and for twice rerating the same statements, were more likely produce the bias than were subjects who were not aware of these experimental demands.

(4) From a motivational perspective, it was predicted that production of the bias would be related to scores on the personality measures. This relationship should have been more apparent within the memory condition than the hypothetical condition, since it was argued that the former is more susceptible to the influence of motivational factors.

For both phases of the experiment, and for the 'no-instruction' group, four dependent variables served as measures of the hindsight bias. For each subject, a rating of a statement without outcome feedback was subtracted from the rating of the same statement made with true-labelled outcome feedback. The average of these difference scores was that subject's true-memory (TM) score. A false-memory (FM) score was computed in the same way but from false-labelled statements. A true-hypothetical (TH) and false-hypothetical (FH) score were similarly calculated, except that the subject rated hypothetical statements only one time and with outcome feedback. The average rating given for the

statement by subjects who had rated the statement without outcome feedback was subtracted from the rating of the statement made by subjects' with outcome feedback. The average of these difference scores was the hypothetical score for that subject. This was either a true-hypothetical (TH) or false-hypothetical (FH) score depending on the form of the outcome feedback, feedback that indicated whether the statement was true or false.

## METHOD

### Subjects

The subjects were 400 undergraduate students enrolled in introductory psychology classes at the University of Manitoba. Subjects were run in groups of approximately thirty, with the experiment requiring a time commitment from each subject of approximately one hour and a quarter. Each subject received two credits toward an experimental participation requirement.

Booklets were randomly assigned to subjects. This resulted in 20 subjects being assigned to each cell in the 2 (Set A or B) by 2 (Direction of Bias) by 2 (Length of Instructions) by 2 (Order) factorial design. An additional 80 subjects were randomly assigned by the same process to a 'no instruction' condition. This resulted in 20 subjects being assigned to each cell in the 2 (Set A or B) by 2 (Order) factorial design.

Experimental Materials

The experimental materials were presented within a single test booklet. The booklet contained a set of 40 general knowledge statements, personality test items, and two final sets of 40 general knowledge statements each. For the entire experiment, 120 statements were rated, the result of 40 of the 80 statements employed being rated twice. These 80 statements were selected from those devised by Hasher, Goldstein and Toppino (1977) and used as the basis for a number of previous investigations of the hindsight bias phenomenon (Campbell & Tesser, 1983; Wood, 1978).

As was the procedure in the Campbell and Tesser (1983) study, the 80 statements were randomly separated into two sets of 40 statements each. Half of the statements within each set were true, half were false. Each set of 40 statements was placed in the first half of the experimental booklets (Appendix B), such that half of the subjects received set A, half set B. For each statement, a 9-point Likert scale, anchored by the labels "certainly false" and "certainly true", was provided.

The next section of the experimental booklet contained the three personality scales. The scale items were presented one after another, with the order of presentation for all booklets being the Marlowe-Crowne Social Desirability scale,

the Texas Social Behavior Inventory, and the Self-Monitoring scale, respectively. The Marlowe-Crowne scale has been found to possess acceptable reliability (KR-20 = .82; Reynolds, 1982), as has the Self-Monitoring scale (KR-20 = .70; Briggs, Check & Buss, 1980). The Texas Social Behavior Inventory has split-half reliability at an acceptable level (r=.97; Helmreich & Stapp, 1974). A five-point Likert scale ranging from "Not at all characteristic of me" to "Very characteristic of me" was provided for responding to each of the items. The scales were responded to on an IBM answer sheet according to the instructions that immediately preceded the entire set of items.

The debiasing instructions, which served as the major experimental manipulation, immediately followed the personality scales. Immediately following the debiasing instructions were forty general knowledge statements, divided into two subsets of twenty statements each. Preceding each subset were brief instructions specific to rating the statements within each subset. Of these two subsets, one was composed of list A statements, one of list B statements. In other words, subjects had previously seen and rated statements from only one of these subsets, while statements from the other subset had not been seen nor rated before.

Additional instructions that directed the subjects to demonstrate their understanding of the bias preceded the rating of two final subsets. Subjects were told to rate the statements within the first of these final two subsets as a subject would who was under the influence of the bias, and the statements within the second of these two final subsets as a subject would who was not under the influence of the bias. Each of these two final subsets contained twenty statements. Ten of the statements within each subset had been rated previously, ten statements had not. Both subsets were immediately preceded by brief instructions specific to rating the statements within each subset.

The post-experimental questionnaire (Appendix C) and the debriefing form (Appendix D) were located at the end of an answer booklet. This answer booklet contained the nine-point scales that were used to record responses to the general knowledge statements. A computer scoring sheet was also provided for subjects to record their responses to the personality measures.

## Procedure

The procedure for the completion of the test booklets was patterned after that of previous research (Campbell & Tesser, 1983) with only those changes necessitated by the experimental manipulations used in this study. An overview

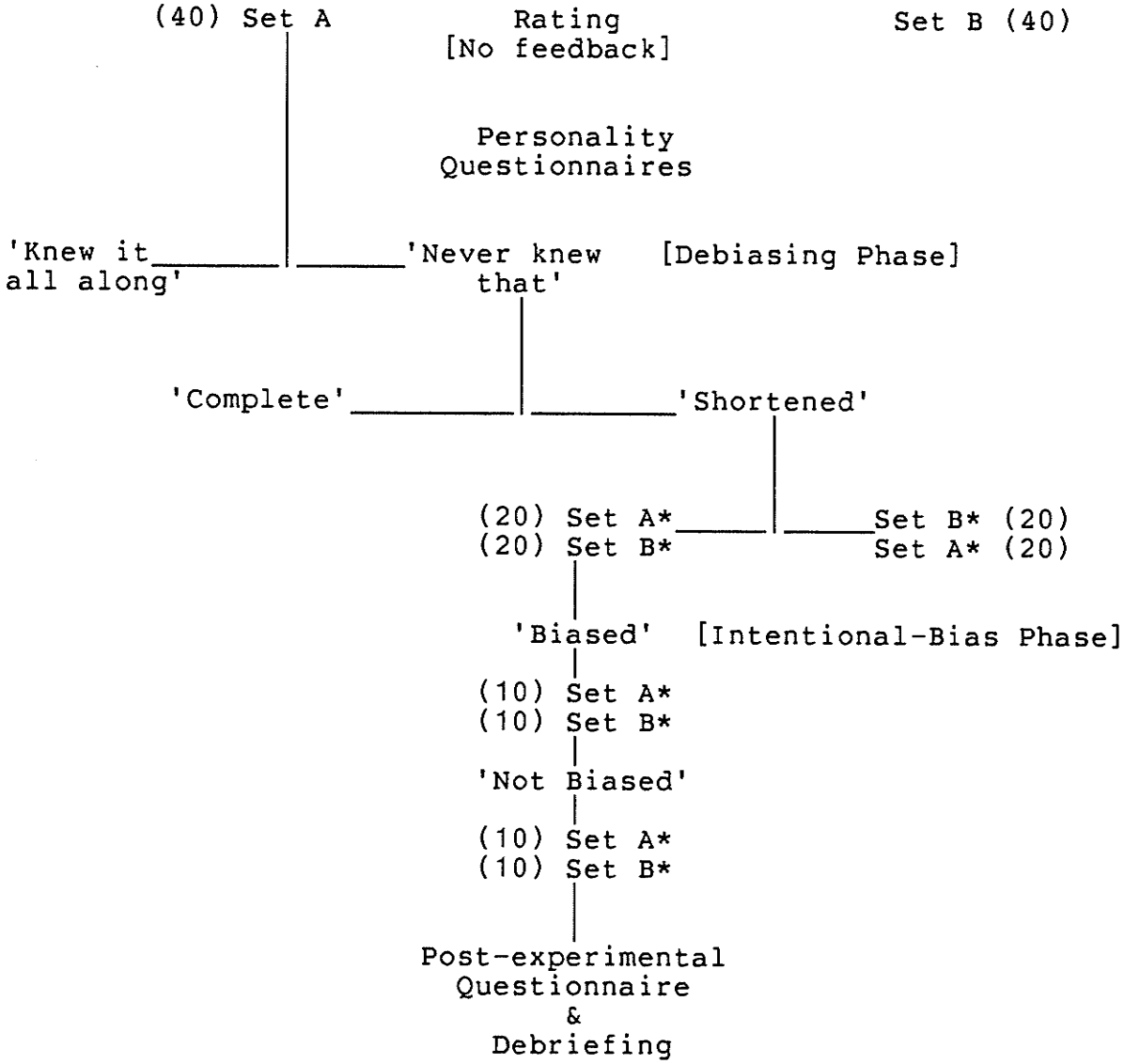
of the procedure that was employed for a subject who received statement set A followed by 'never-knew-that'/'shortened' instructions is provided in Table 1. Comparable procedures were followed by subjects who initially received statement set B or other forms of the debiasing instructions.

Upon arrival at the classroom in which the experiment took place, each subject was given an experimental booklet and was instructed not to open it until after the experimenter provided verbal instructions. The following instructions were read by the experimenter.

> Thank you for agreeing to participate in this experiment. The experiment will take approximately one hour and a quarter, and is worth 2 experimental credits.
>
> All of the instructions you will need will be found in the experimental booklets I have given you. It is very important you follow these instructions exactly. Do not speak with other participants and do not write in the booklets. You will record all your answers in an answer booklet or on a computer scoring sheet, both of which are provided.

Table 1

Overview of the Experimental Procedure




* Outcome feedback provided
Values in brackets are the number of statements.
Half of each set of statements are true, half are false.

> When I say "begin", please carefully read the instructions on the first two pages of the booklet. Complete the first 40 statements by marking your responses in the answer booklet.
>
> When you have completed the 40 statements, do not go on. Leave the booklets open and wait for further instructions. You may now begin reading the instructions on the first two pages of the experimental booklet.

Written instructions taken from Campbell and Tesser (1983) were also read by the subjects. These instructions are reproduced on the first two pages of the experimental booklet (Appendix B).

Once all of the subjects completed the first set of statements, the following instructions regarding the personality measures were read by the experimenter.

> Now turn to the next set of instructions. Those instructions will tell you how to respond to the next 83 items. Read the instructions carefully. Do not use the answer booklet you just employed to rate the general knowledge statements. The computer scoring sheet is to be used with this section. When you have completed these items, please wait for further instructions. Do not go

> on or go back to an earlier section of the experimental booklet. Leave the experimental booklet and the answer booklet open, your pencil down, and wait for further instructions.

Subjects were allowed fifteen minutes to complete the personality measures. Fifteen minutes was found to be sufficient time for all subjects to complete the personality items. The following set of instructions were then read to the subjects prior to completion of the remainder of the experimental booklet.

> Now, please turn to the next set of instructions. Complete the remainder of the experimental booklet using the answer booklet to record your responses. Please read the instructions you are provided carefully. Do not return to an earlier section of the booklet. When the booklets have been completed, you will find a final short questionnaire at the end of the answer booklet. Please complete the questionnaire, and then read the explanation for the study that follows. You are welcome to remove the explanation and take it with you.

Subjects then read one of four forms of the instructions that manipulated the experimental demands (Appendix B).

That is, subjects read either 'complete'/'never-knew-that', 'complete'/'knew-it-all-along', 'shortened'/'never-knew-that', or 'shortened'/'knew-it-all-along' versions of the debiasing instructions, or 'no instructions' were provided.

'Complete' instructions were patterned after those given to the debiasing group in the Fischhoff (1977) study, modified for use with materials employed in the present study. 'Shortened' instructions were those same instructions except that the last two paragraphs were not included. Within the 'shortened' and 'complete' instructions, the 'knew-it-all-along' instructions were the modified descriptors of the hindsight bias provided by Fischhoff (1977). The 'never-knew-that' instructions were constructed to give a description of the hindsight bias opposite to that which was given in the 'knew-it-all-along' condition. Words in parenthesis in the experimental booklet (Appendix B) indicate the variations on the instructions for the 'knew-it-all-along' and the 'complete' conditions.

Two subsets of twenty general knowledge statements each were then completed. Statements within both subsets, and for all subsequent statement subsets, were accompanied by feedback that indicated whether the statement was true or false. Half of the statements within each subset were true, half were false. Subjects were instructed to ignore this

feedback while completing their ratings. One of these subsets was made up of general knowledge statements that had been rated within the initial set of 40 statements. Subjects were instructed to assign the same value to each statement as they had the first time they had rated the statement. This was termed the 'memory' condition. The second subset was made up of general knowledge statements that the subjects had not rated previously. Subjects were instructed to assign the rating that they would have given the statement if they had not been provided with the correct answer. This was termed the 'hypothetical' condition. Half of the subjects received the 'hypothetical' subset first, half received the 'memory' subset first.

Prior to rating two final subsets, 'biased' and 'not biased' instructions were provided to subjects. Subjects were asked to demonstrate their understanding of the bias by rating the first of the final two subsets in a 'biased' manner, i.e., under the influence of the 'knew-it-all-along' or 'never-knew-that' effect. The second of these two subsets was to be rated in a non-biased manner, i.e. as they should if they were not under the influence of the 'knew-it-all-along' or 'never-knew-that' effect. Half of the statements within each subset were memory statements, half hypothetical statements, with the order for memory and hypothetical statements the same as for those subsets rated

immediately after the personality measures. Additional brief instructions preceded each subset.

After each subject completed the final two subsets, a post-experimental questionnaire (Appendix C) was completed. The subject then read the debriefing, presented in Appendix D, which included an explanation of the purpose of the research, a discussion of the deception involved, a refererence to prior research on the topic, and assurance of anonymity.

# RESULTS

## Debiasing Phase

Analysis of the first phase of the study, i.e. the subject's confidence ratings after having received feedback, was conducted by a 2x2x2x2 between-subjects MANOVA. The four independent variables were Set (List A or List B), Order ('memory then hypothetical' or 'hypothetical then memory'), Direction ('knew-it-all-along' or 'never-knew-that'), and Length ('complete' or 'shortened'). Univariate ANOVAs and Scheffes were employed for post-hoc comparisons. Assessments of whether the mean ratings differed significantly from zero were made through *t* tests. Missing ratings were replaced by the mean rating appropriate for the statement and for the experimental condition to which the subject had been assigned. With 20 subjects per cell, multivariate normality was assured (Tabachnick & Fidell, 1983, p. 232). Homogeneity of variance was implied by the equal sample sizes, and was supported by Bartlett's test. Cook's *d* statistic (Cook, 1979) revealed no outliers within any of the four dependent variables. Ploting the dependent variables against each other did not reveal any significant departures from linearity.

Prior to examining the extent to which the manipulations impacted on the hindsight bias, one must first demonstrate that the hindsight bias was present. Operating under the null hypothesis for treatment effects, the four hindsight bias measures were considered collapsed across the experimental manipulations. Replicating past research, a hindsight bias was found with true-labelled and false-labelled hypothetical statements (TH and FH) and false-labelled memory statement (FM, Table 2). A hindsight bias was not found with true-labelled memory statements (TM). As with the dependent variables for which the bias had been found, the mean rating for TM was in the expected direction. Ratings of hypothetical statements were significantly greater than ratings for memory statements, both for true-labelled ( $t$ (319)= 7.04, $p$ <.0001) and false-labelled ( $t$ (319)= -9.56, $p$ <.0001) statements.

The same set of data was then examined for the impact of the debiasing manipulations on the hindsight bias. The results from the MANOVA that tested the effects of these manipulations are presented in Table 3. Contrary to the first hypothesis, the debiasing manipulations of Direction and Length did not significantly influence the hindsight bias. The only significant multivariate main effect was for Set ( $F$ (4,301)= 3.95, $p$ <.0039), a somewhat unexpected finding since the two sets of statements were randomly

Table 2

Summary Statistics for the Four Dependent Variables, Collapsed Across the Experimental Manipulations, Debiasing Phase

| | M | Var | t | p |
|---|---|---|---|---|
| True Memory | .04 | .25 | 1.56 | .1198 |
| False Memory | -.22 | .37 | -6.54 | .0001 |
| True Hypothetical | .37 | .53 | 8.98 | .0001 |
| False Hypothetical | -.71 | .50 | -17.88 | .0001 |

Note: degrees of freedom for each t-test were 319.

Table 3

MANOVA Summary Table for the Debiasing Phase

| Variable(s) | F | p |
|---|---|---|
| Set (S) | 3.95 | .0039 |
| Direction (D) | .77 | .5481 |
| Order (O) | 1.33 | .2600 |
| Length (L) | .16 | .9593 |
| SxD | 2.36 | .0539 |
| SxL | .48 | .7498 |
| SxO | .41 | .8003 |
| DxL | .41 | .8027 |
| DxO | .43 | .7870 |
| LxO | .37 | .8317 |
| SxDxL | .43 | .7867 |
| SxDxO | .88 | .4771 |
| SxLxO | .70 | .5952 |
| LxOxD | .88 | .4755 |
| SxDxLxO | .95 | .4330 |

Note: degrees of freedom for each test were 4 and 301.

created. Univariate analyses (Appendix E) revealed a significant effect of Set for FM. For FM, the mean for List B was -.33, for List A -.11 ( $p < .05$). No other multivariate or univariate main effects or interactions were statistically significant. Cell means for main effects are presented in Table 4, and for interaction effects, in Appendix E.

Table 4

Cell Means for the Four Measures of the Hindsight Bias, Debiasing Phase

| Variables | Hypothetical | | Memory | |
|---|---|---|---|---|
| | True | False | True | False |
| Set | | | | |
| A | .29 | -.75 | .07 | -.11 |
| B | .44 | -.67 | .02 | -.33 |
| Direction | | | | |
| Knew | .36 | -.67 | .01 | -.30 |
| Never | .38 | -.75 | .07 | -.24 |
| Length | | | | |
| Complete | .34 | -.69 | .04 | -.21 |
| Shortened | .39 | -.73 | .04 | -.23 |
| Order | | | | |
| Hypo/Mem | .42 | -.76 | .06 | -.28 |
| Mem/Hypo | .32 | -.66 | .03 | -.17 |

Intentional Bias Phase

The second phase of the experiment involved examination of responses on the hindsight bias measures while subjects were operating under 'biased' or 'not biased' instructions. A 2x2x2x2x2 mixed MANOVA served this purpose. The additional factor was the within-subject factor Understanding, this being whether statements were rated under 'biased' or 'not biased' instructions, respectively. Univariate ANOVAs and Scheffes were again employed for post-hoc comparisons, as were *t* tests to assess whether hindsight bias scores differed significantly from zero. Missing ratings were handled as in phase one. Twenty subjects per cell again assured robustness to non-normality and variance homogeneity, the latter supported by Bartlett's test. Cook's *d* statistic revealed no outliers for any of the four dependent variables, and examination of the dependent variables plotted against each other did not reveal any gross departures from linearity.

The Intentional-Bias Phase was run to assess the extent instructions to respond in a biased or a non-biased way would impact on the hindsight bias measures and would interact with the experimental manipulations from the Debiasing Phase. The second hypothesis predicted that subjects would provide biased or non-biased ratings when the demands for each were made explicit, and that these

demonstrations would be consistent with the explanation that had been provided for the bias within the Debiasing Phase. This hypothesis was supported. The results from the MANOVA that assessed the impact of the Understanding factor within the debiasing conditions are presented in Table 5. A significant multivariate main effect was found for the Understanding factor ( $F$ (4,301)= 18.83, $p$ <.0001). With the exception of FH, significant univariates were found with all four measures of the hindsight bias (Appendix F). Cell means for the Understanding variable are presented in Table 6. 'Biased' instructions resulted in a much stronger bias than 'not biased' instructions, a difference most apparent with memory scores.

A significant multivariate main effect was also found for Direction ( $F$ (4,301)= 26.56, $p$ <.0001). The univariates revealed that Direction impacted on each of the four measures of the bias. Examination of cell means (Table 7) reveals a much stronger bias for the 'knew-it-all-along' over the 'never-knew-that' condition, especially for the hypothetical statements.

The main effects of Direction and Understanding are interpretable in terms of their significant multivariate interaction ( $F$ (4,301)= 31.71, $p$ <.0001). All corresponding univariates were significant. Examining the cell means for TH and FH (Figure 1) reveals that subjects responding under

Table 5

MANOVA Summary Table for the Intentional-Bias Phase

| Variable(s) | F | p |
|---|---|---|
| Set (S) | 3.35 | .0106 |
| Direction (D) | 26.56 | .0001 |
| Order (O) | 4.07 | .0032 |
| Length (L) | 1.49 | .2063 |
| Understanding (U) | 18.83 | .0001 |
| SxD | 1.79 | .1313 |
| SxL | 2.24 | .0649 |
| SxO | 5.88 | .0001 |
| DxL | 2.42 | .0488 |
| DxO | 1.07 | .3706 |
| LxO | 2.71 | .0305 |
| UxS | 6.68 | .0001 |
| UxD | 31.71 | .0001 |
| UxL | .58 | .6808 |
| UxO | 10.27 | .0001 |
| SxDxL | 1.14 | .3365 |
| SxDxO | 1.44 | .2220 |
| SxLxO | .78 | .5392 |
| LxOxD | .60 | .6665 |
| SxDxU | 2.39 | .0510 |
| SxLxU | 2.69 | .0315 |

| | | |
|---|---|---|
| SxOxU | 6.06 | .0001 |
| LxOxU | 1.94 | .1030 |
| DxLxU | .56 | .6890 |
| DxOxU | .58 | .6770 |
| SxDxLxO | 3.26 | .0012 |
| SxDxLxU | .58 | .6756 |
| SxDxOxU | .66 | .6212 |
| SxLxOxU | 1.04 | .3846 |
| DxLxOxU | .31 | .8702 |
| SxDxLxOxU | .93 | .4480 |

Note: degrees of freedom for each test were 4 and 301.

'knew-it-all-along' instructions produced a stronger hindsight bias with 'biased' instructions ( $p$ <.001). Subjects responding under 'never-knew-that' instructions, in contrast, produced a stronger hindsight bias with 'not biased' instructions ( $p$ <.05 for TH, $p$ <.001 for FH). The same pattern of results were found for TM and FM (Figure 2), except that only for the 'knew-it-all-along' condition was the Understanding variable significant ( $p$ <.001).

A number of other significant main effects and interactions were found (Appendix G). These findings did not impact on the main effects of Direction and Understanding, or the Direction by Understanding interaction. In addition, these findings were not found with more than two of the four hindsight bias measures. Finally, these findings were not predicted by the results from the Debiasing Phase.

Table 8 indicates that, collapsing across experimental manipulations in the Intentional-Bias phase, the hindsight bias was again replicated. All four dependent measures showed a strong hindsight bias, with means in the predicted direction. Once again, hypothetical ratings were larger in magnitude than memory ratings for both true-labelled ( $t$ (319)= 5.24, $p$ <.0001) and false-labelled ( $t$ (319)= -4.36, $p$ <.0001) statements.

Table 6

Mean Values for the Understanding Variable, Intentional-Bias Phase

| | Biased | Not Biased |
|---|---|---|
| True Hypothetical | .74[a] | .48[a] |
| False Hypothetical | -.73[a] | -.78[a] |
| True Memory | .65[a] | .02 |
| False Memory | -.73[a] | -.29 |

Note: a = Significantly different from zero, $p<.0001$

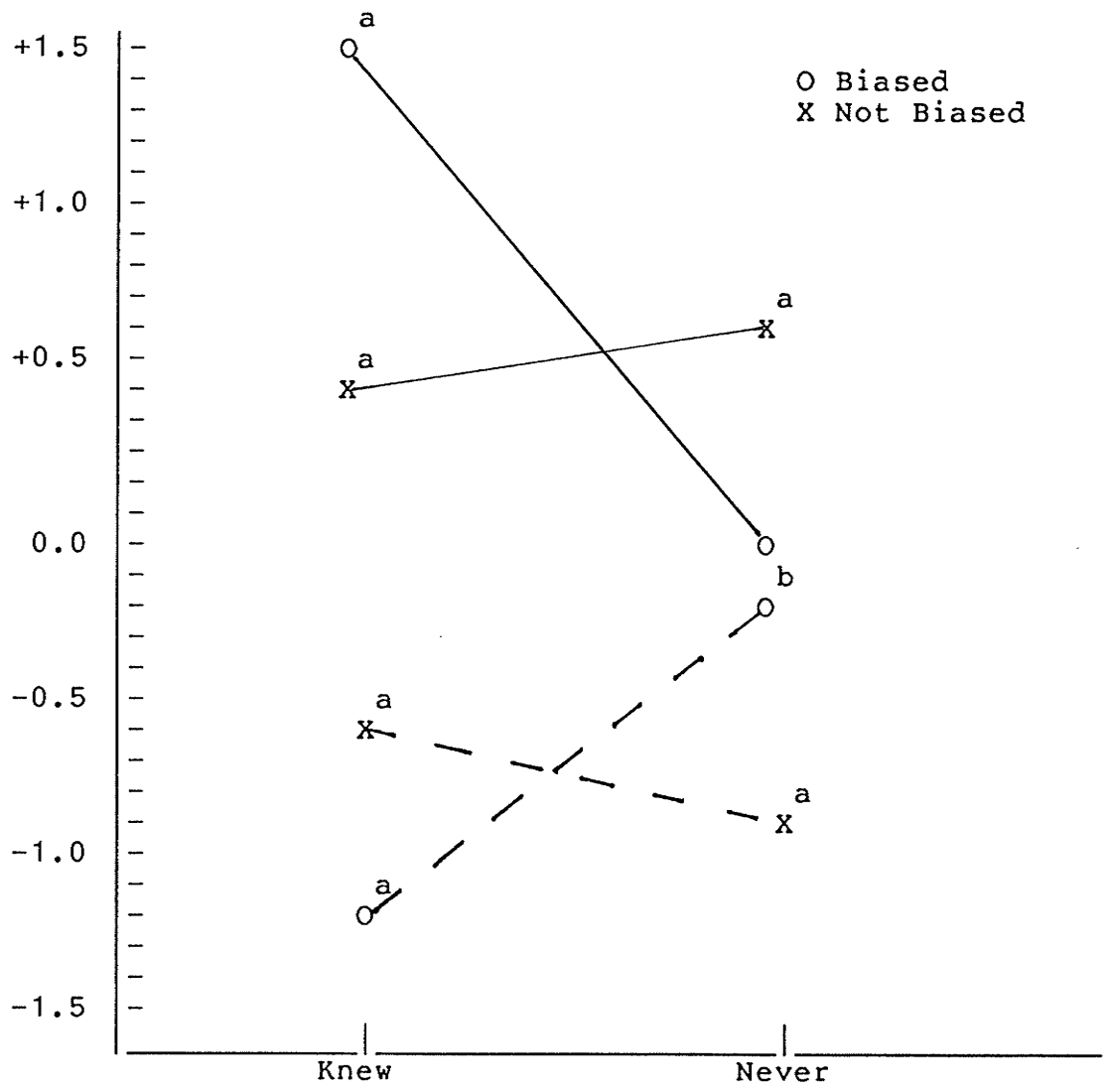
Table 7

Mean Values for the Variable Direction, Intentional-Bias Phase

| | Knew | Never |
|---|---|---|
| True Hypothetical | .92[a] | .30[a] |
| False Hypothetical | -.93[a] | -.57[a] |
| True Memory | .72[a] | -.06 |
| False Memory | -.88[a] | -.15[b] |

Note: a = Significantly different from zero, $p<.0001$
b = Significantly different from zero, $p<.05$

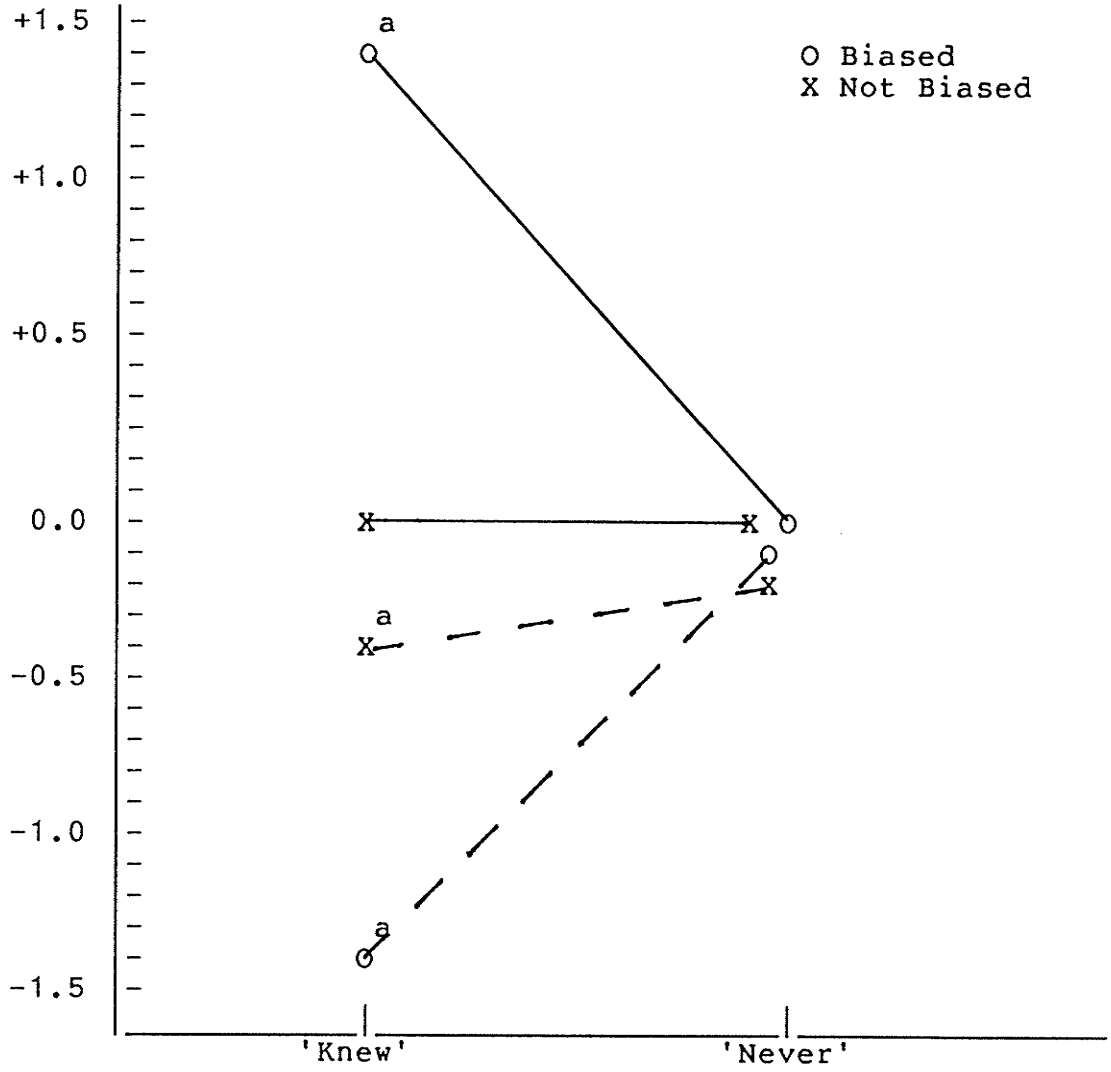
Figure 1: Plot of the Direction by Understanding interaction, Dependent Variables True and False Hypothetical, Intentional-Bias Phase




Note: a = significantly different from zero, $p<.0001$
b = significantly different from zero, $p<.05$

_____ True-Hypothetical Scores
_ _ _ False-Hypothetical Scores

Figure 2: Plot of the Direction by Understanding Interaction, Dependent Variables True and False Memory, Intentional-Bias Phase




Note: a = significantly different from zero, $p<.0001$

_____ True-Memory Scores
_ _ _ False-Memory Scores

Table 8

Summary of the Four Dependent Variables, Collapsed Across the Experimental Manipulations, Intentional-Bias Phase

| | M | Var | t | p |
|---|---|---|---|---|
| True Hypothetical | .61 | 1.51 | 12.60 | .0001 |
| False Hypothetical | -.75 | 1.49 | -15.64 | .0001 |
| True Memory | .33 | 1.82 | 6.28 | .0001 |
| False Memory | -.51 | 1.99 | -9.17 | .0001 |

Note: degrees of freedom for each test were 319.

No Instruction Condition

Control subjects were examined for replication of the hindsight bias, for awareness as to the purpose of the outcome feedback, and for the role of personality factors in the bias. The personality scores for a subsample that replicated Fischhoff's original debiasing condition, and the post-experimental questionnaire responses for a subsample that were balanced for form of debiasing instructions, were examined solely in terms of comparison with the responses of the 'no instruction' subjects. The results of these comparisons are presented in this section.

While the 'no-instruction' group could have been integrated into the factorial phases of the study (Himmelfarb, 1975), the data from this sample were examined separately to facilitate analysis and to be congruent with the hypotheses. The ratings of statements without outcome feedback by subjects in the factorial portion of the design were used to obtain the hypothetical scores. Ratings from this larger sample were predicted apriori to be more stable than those provided by the smaller 'no-instruction' sample. Missing hindsight bias ratings were replaced by the appropriate mean values. Missing post-experimental questionnaire and personality measure values were not replaced. Since at least 20 subjects per cell were available for all analyses performed, the assumption of normality for

the hindsight bias measures was presumed to have been met. Cook's d statistic revealed four outliers within the hindsight bias measures. Since their inclusion did not significantly alter the results, these scores were not deleted. Examination of the plots between the hindsight bias measures revealed linear interrelationships.

Table 9 reveals that, for the 'no-instruction' sample, a hindsight bias was found for both true and false hypothetical scores. No hindsight bias was found with memory scores, although all four dependent measures reflect mean scores that differed from zero in the direction predicted for the hindsight bias. As before, hypothetical scores exceeded memory scores, both for true-labelled ( $t$ (79)= 4.01, $p$ <.0001) and for false-labelled ( $t$ (79)= -4.76, $p$ <.0001) statements.

Post-experimental Questionnaire. The author was blind to subjects' hindsight bias scores while scoring the the post-experimental questionnaires. Summary statistics for the responses on the post-experimental questionnaire are presented in parentheses in Appendix C. Responses from the subsample drawn from the debiasing phase of the study were found to be similar to those provided by 'no-instruction' subjects. Of the 'no instruction' subjects, more than a third were capable of communicating what the experimenter expected to find in the study (N=33 for 'no instruction'

Table 9

Summary Statistics for the No Instruction Group

| | M | Var | t | p |
|---|---|---|---|---|
| True Memory | .02 | .32 | .28 | .7822 |
| False Memory | -.06 | .27 | -1.04 | .3038 |
| True Hypothetical | .43 | .61 | 4.88 | .0001 |
| False Hypothetical | -.69 | .92 | -6.24 | .0001 |

Note: degrees of freedom for each test were 79.

subjects, N=34 for the Debiasing Phase subsample) and how subjects were expected to rate the second set of statements (N=35, 23). More than a half offered an appropriate explanation as to why the same statements were rated twice (N=50, 56). Judgements as to how ratings of the second set of statements were to be made, both for true ( $\underline{M}$ =4.84, 5.00) and for false ( $\underline{M}$ =5.01, 4.72) statements, fell between 'closer to the extreme' and 'closer to the middle' of the nine-point scale. Difficulty in rating statements for which outcome feedback had been provided was only moderate ( $\underline{M}$ =5.59, 6.08), although a large proportion of subjects recognized that the outcome feedback was expected to influence ratings (N=55, 63), stated that the outcome feedback did have some effect on their ratings (N=59, 58), and admitted that some attempt was made to avoid the bias (N=57, 58).

An overall 'awareness' measure was constructed. Those subjects who offered what was judged to be an appropriate explanation for why the same statements were rated twice (N=50, 56) were classified as 'aware'. The third hypothesis was that those subjects who responded in a biased manner could be differentiated from those subjects who did not according to awareness as to the the purpose of the outcome feedback. Four separate analyses of variance were employed, with the four measures of the hindsight bias as the

dependent variables, and the awareness score as the independent variable. The absence of significant main effects for awareness for the 'no instruction' sample failed to support the third hypothesis (Table 10). It was thus not possible to differentiate those subjects who showed a bias from those subjects who did not in terms of the awareness measure. Cell means for the awareness variable are presented in Table 11, and again show the superiority of the bias with hypothetical scores. There was a significant main effect for awareness within the Debiasing Phase sample (Table 12). Contrary to the expectation that experimental awareness would be associated with production of the hindsight bias, those subjects who were classified as aware produced a significantly smaller bias for memory statements compared to subjects who were classified as unaware. Cell means are presented in Table 13.

Personality Measures. The fourth hypothesis was that, for the 'no instruction' sample, demonstration of the bias would be related to scores on the personality measures. To test this hypothesis, responses on the Texas Social Behavior Inventory, the Self-Monitoring scale and the Marlowe-Crowne scale were predictor variables, and the four measures of the hindsight bias criterion variables, in a multiple regression analysis. A stepwise backward selection procedure (Tabachnick & Fidell, 1983) was employed to assess which

Table 10

Summary of ANOVAs for the Awareness Variable, No Instruction Group

| Source | df | SS | F | p |
|---|---|---|---|---|
| True Hypothetical | | | | |
| Model | 1 | .00 | 0.00 | .9505 |
| Error | 78 | 48.41 | | |
| False Hypothetical | | | | |
| Model | 1 | .33 | 0.36 | .5525 |
| Error | 78 | 72.32 | | |
| True Memory | | | | |
| Model | 1 | .00 | 0.00 | .9630 |
| Error | 77 | 25.13 | | |
| False Memory | | | | |
| Model | 1 | .36 | 1.32 | .2543 |
| Error | 77 | 20.87 | | |

Table 11

Summary of the Mean Awareness Scores
for the No Instruction Group

| | Aware N=50 | Unaware N=29 |
|---|---|---|
| True Hypothetical | .43 [b] | .42 [b] |
| False Hypothetical | -.62 [a] | -.75 [b] |
| True Memory | .02 | .02 |
| False Memory | -.11 | .02 |

Note a = Significantly different from zero, $p < .0001$
b = Significantly different from zero, $p < .001$

Table 12

Summary of ANOVA's on Awareness for the Debiasing Subsample

| Source | df | SS | F | p |
|---|---|---|---|---|
| True Hypothetical | | | | |
| Model | 1 | .64 | 0.96 | .3297 |
| Error | 79 | 52.89 | | |
| False Hypothetical | | | | |
| Model | 1 | .31 | 0.51 | .4764 |
| Error | 79 | 47.93 | | |
| True Memory | | | | |
| Model | 1 | 3.38 | 11.15 | .0013 |
| Error | 79 | 23.95 | | |
| False Memory | | | | |
| Model | 1 | 2.12 | 4.98 | .0285 |
| Error | 79 | 33.62 | | |

Table 13

Summary of the Mean Awareness Scores
for the Debiasing Phase Subsample

| | Aware<br>N=56 | Unaware<br>N=24 |
|---|---|---|
| True Hypothetical | .38 | .57[a] |
| False Hypothetical | -.72[a] | -.85[a] |
| True Memory | -.11 | .32 |
| False Memory | -.22 | -.57[a] |

a = Significantly different from zero, $p < .0001$

predictor variables, if any, were appropriate for inclusion. For the 'no instruction' sample, step 2 of the regression for TH revealed that the Marlowe-Crowne scale was the only significant predictor (Appendix H, Table I-1). High social desirability was found to predict biased responding. Table H-2 indicates that for FH, the best model after step 2 was a one variable model, that variable being the Texas Social Behavior Inventory. Low self-esteem was found to predict biased responding. Tables H-3 and H-4 reflect that for TM and FM, none of the personality measures, separately or in combination, were significant predictors. Table H-5 presents, for the 'no instruction' sample, summary statistics for the personality measures.

The responses on the personality measures of subjects who replicated Fischhoff's original debiasing condition (i.e. 'knew-it-all-along'/complete instructions) were compared with their ratings of the hindsight bias statements. Table H-6 indicates that for FH, the best model after step 2 was a one variable model, that variable being the Texas Social Behavior Inventory. This finding failed to attain statistical significance. Contrary to the findings with the 'no instruction' sample, high self-esteem was found to predict biased responding. For TH, TM and FM, none of the personality measures, separately or in combination, were significant predictors (Tables H-7 to H-9). Table H-10

presents summary statistics for the three personality measures.

# DISCUSSION

## Replication of the Hindsight Bias

Prior to examining the manipulations of the hindsight bias, it seemed essential to demonstrate that the classic hindsight bias pattern of results was replicated. The present study found the hindsight bias to once again be a 'robust' effect. Each of the four dependent measures were evaluated within each of the two phases and the 'no instruction' condition for demonstration of the hindsight bias. A significant hindsight bias was found for nine of these twelve assessments, suggesting continuity between the present study and past research. In the three cases in which the scores did not significantly differ from zero, means were in the predicted direction. The size of the hindsight bias was impressive. The difference from the first to the second set of ratings was approximately one-half of one point on the nine-point scale.

Two factors may have operated to make even this value an underestimation of the strength of the bias. First, if a subject was certain that a statement was true or false on first rating, and had assigned a value that reflected this

extreme confidence, then subsequent confirmation of this judgment by outcome feedback could not result in a higher second rating for the statement. For instance, if a subject had rated a statement as +4, extreme confidence that the statement was true, then the rating given after confirmatory evidence could be no higher than +4, with the result being a difference score of zero. Second, it was also unlikely than an extreme rating for a statement would be disconfirmed by the outcome feedback. The difficulty level of most statements would cause their first ratings to cluster around the midpoint of the scale. On these items, a difference score of one half a point should not be judged in terms of the nine point scale, but in terms of a smaller segment of the total scale.

Another consistent finding from the hindsight bias literature that was evident in this study was the superiority of hypothetical scores. This was apparent for both true-labelled and false-labelled statements within both phases of the experiment and with the 'no treatment' control group. The superiority of the hypothetical procedure over the memory procedure for obtaining a larger hindsight bias was noted by Campbell and Tesser (1983), who attributed it, in part, to the 'between' versus 'within' approaches used to obtain each score. A larger bias for hypothetical statements is not incongruent with the cognitive, the motivational or

the demand characteristic explanation for the effect. From the cognitive perspective, the hypothetical situation provides less information to be integrated within the store of available knowledge. The only rating of a hypothetical statement that a subject has available is that provided by the outcome feedback. From the motivational perspective, responding within a memory format might be seen as invoking a factor that demands congruity between the two sets of ratings. From the demand characteristic perspective, a subject responding within a memory format might recognize that he or she is expected to adjust the second set of ratings in the direction of the outcome feedback but is constrained in this adjustment by the first set of ratings given.

## Effect of the Debiasing Manipulations

While contrary to the first hypothesis, the lack of significant results within the Debiasing Phase was not inconsistent with past research. Providing instructions for the bias that excluded the final passage in Fischhoff's debiasing instructions, a passage judged to be contrary to the preceding debiasing instructions, was expected to have some impact. It did not. More importantly, Fischhoff's demonstration that debiasing instructions were not effective in preventing the bias from occuring was replicated in the

failure to find an effect of Direction within the Debiasing Phase. This support must be considered tentative since, as with Fischhoff's original demonstration, a finding of no effect is ambiguous. One alternative to Fischhoff's explanation for the lack of debiasing efficacy is the nature of the debiasing instructions. Fischhoff presumed that these were potentially effective instructions. In other words, it was presumed that these debiasing instructions would be effective with a less 'robust' bias or with a bias that was the product of an artifact. These instructions may instead be ineffective by their nature, rather than by the nature of the phenomenon, an ineffectiveness that may not have been remedied by the minor changes made within the present study.

Effect of the Intentional Bias Manipulation

The Intentional Bias Phase was structurally a replication of the Debiasing Phase with the addition of the Understanding factor. It was predicted that ratings given under 'biased' instructions would be more in the direction of the outcome feedback than ratings given under 'not biased' instructions, and that the magnitude of this effect would vary according to whether a 'knew-it-all-along' or a 'never-knew-that' form of an explanation for the bias was provided. For all of the measures of the dependent variable except FH, a larger hindsight bias was found with 'biased'

instructions. The 'not-biased' instructions significantly diminished the bias scores, compared to the effect of 'biased' instructions and to scores collapsed across all treatment conditions (i.e. the values used to demonstrate a hindsight bias). A larger bias was found within the 'knew-it-all-along' condition than within the 'never-knew-that' condition. The 'never-knew-that' condition, however, did not produce the hypothesized 'reverse' bias. Values were closer to, but did not cross zero, as would be demanded for a 'never-knew-that' bias. Nevertheless, the manipulation of the instructions apparently did have some effect.

An explanation for these two main effects can be derived from the significant interaction of Direction by Understanding, found for both true and false-labelled memory and hypothetical scores. For 'knew-it-all-along' subjects, both for memory and hypothetical statements, the instructions to rate statements in a more or in a less biased way were complied with in a manner that was according to their understanding of the bias. These subjects were capable of producing a powerful bias or a significantly reduced bias when the demand to do so was made explicit. For 'never-knew-that' subjects, and for memory statements, the 'never-knew-that' effect was not found. Their ratings did not differ whether under biased or not biased

instructions, and reflected a significant hindsight bias. Perhaps the explanation for going against the outcome feedback was not sufficiently compelling. Alternatively, memory ratings were constrained by the subject having previously committed himself to a judgement for the memory statements. With hypothetical statements, 'never-knew-that' subjects were not constrained by past ratings. Instead, these subjects were free to operate according to their understanding of the bias. 'Biased' rating scores did depart from the direction suggested by the outcome feedback, falling below the level found with 'not biased' instructions. This movement was not sufficient such that a reverse hindsight bias was found, or even such that a significant hindsight bias was not found. Nevertheless, the 'never-knew-that' instructions had impacted on ratings in the predicted manner.

From the strict cognitive perspective, the hindsight bias should not have been influenced by either the Direction or Understanding manipulations. In neither instance was the outcome feedback discredited or in any way altered. Instead, making the demands for unbiased responding more explicit did result in a diminished bias, a bias that was influenced by the form of explanation given, either for a 'knew-it-all-along' or a 'never-knew-that' effect

Post-experimental Questionnaire Responses

A large number of the 'no instructions' subjects recognized that the function of the hindsight bias paradigm was to assess whether outcome feedback would influence their subsequent ratings. A number of subjects also initially described the study as a test of their honesty, which might serve to explain large differences between the hypothetical and memory produced scores. With memory statements, there may be an implicit demand to rate the statements as one had previously, conflicting with experimental demands to submit to the influence of the outcome feedback. With hypothetical statements, a subjects might convince himself that he had rated the statements without the benefit of the outcome feedback.

No differences on the hindsight bias measures were found for 'no instruction' subjects classified as 'aware' or 'unaware', and only with Debiasing subjects for memory statements. The 'pact of ignorance' notion (Orne, 1962) suggests that a proportion of subjects might have been aware but unwilling to provide post-experimental questionnaire responses that would have permitted their classification as 'aware'. If one conjectures that virtually all subjects were aware, rather than the half of the sample so classified, then the difference between the hindsight bias scores of those who were detected as being aware would not be found to differ from those who were not detected.

The finding that aware subjects produced less of a bias with memory statements seems, at first, to be contradictory to predictions made. Furthermore, this effect was found with memory statements, while a much stronger bias was consistently found with hypothetical statements. Subjects who recognized that the ratings of the second set of statements were to be influenced by having been provided with the correct answer did not differ from those who did not possess this understanding on hypothetical statements. Ratings of memory statements were constrained by past ratings. It might be hypothesized that aware subjects were more cognizant of being constrained by past ratings, and choose to rate only hypothetical items in a biased manner. Support for this position is found in the consistently greater bias found with hypothetical statements for much of the sample, the majority of whom were classified aware.

## Motivational Variables

The personality measures provided some limited support for the motivational perspective. For 'no instruction' subjects and true and false hypothetical statements, a personality measure was found to be a significant predictor of biased responding. More telling evidence within the limits of the correlational design would have been more personality measures as significant predictors of biased

responding, larger effects for these indicies in explaining the variance within the hindsight measures, and consistency as to the personality measures that acted upon the bias. No bias was found for memory scores, contrary to the prediction, making it difficult to evaluate the consequence of this finding for the personality explanation.

## Conclusions and Implications

This study did nothing to dispel the notion that the laboratory demonstration of the bias is a robust effect. The bias has been demonstrated in a number of contexts and with a variety of experimental materials. What was of interest was the source of this bias. This study focused on the bias within the laboratory context, and argued that production of the bias, within this context, is the product of demand characteristics. Support for this position was found with the demonstration that production of the bias could be influenced by instructions provided to subjects. Specifically, it was found that instructions to rate statements in a biased or a non-biased manner were effective, and that different forms of an explanation for the bias impacted on ratings under such instructions. Fischhoff attempted the latter, and having found no evidence for the instructions provided to subjects as influencing their decidedly biased responding, concluded that demand

characteristics were not involved in production of the bias. The present study suggests that it is possible to provide experimental cues to subjects that will alter the size and the direction of the bias.

One implication of accepting demand characteristics as a significant factor in the production of the bias is that the validity of past research that has employed the experimental paradigm is cast in doubt. If the research has studied what amounts to an artifact of experimental participation, then conclusions drawn, while internally valid, would have no relation to behavior outside of the laboratory. For instance, methods of ameliorating the bias that have developed from the laboratory paradigm may have no practical application. If one accepts that the laboratory produced bias is an artifact, then the issue of the existance of the bias beyond the confines of the laboratory must also be questioned. This concern is significiant since past examinations of the bias, distinct from the laboratory context, have relied upon the laboratory paradigm. Furthermore, subjects in these studies often have been aware that they were participating in an experiment. This demand characteristic interpretation suggests a much less robust phenomenon, a notion consistent with the weaker level of the bias found in non-laboratory contexts in which experimental demands are less often present.

Knowledge of the outcome of significant events does distort attitudes and behavior. The present demand characteristic analysis challenges the laboratory demonstration of the phenomenon, not the issue of whether the bias exists. The present study suggests that researchers lack an experimental paradigm that permits the demonstration, manipulation, and study of the non-artifactual hindsight bias in a laboratory context. There are a number of responses to this conclusion. One alternative is to accept that the bias does exist distinct from the laboratory, and that the bias within the current paradigm is the combination of both psychological and artifactual processes. The challenge is to then devise a means for separating the two processes. One possibility is the development of a sophisticated post-experimental questionnaire. With such an instrument, a researcher could determine which subjects are being unduely influenced by participation-related factors. The responses from 'aware' and 'non-aware' subjects could then be compared. Although an attempt along these lines was made in the present study, the post-experimental questionnaire employed was probably not sufficiently sensitive for this task. Alternatively, an experimental paradigm could be developed in which artifactual demands could be demonstrated to be of less impact than with the present paradigm, or circumstances could be sought in which demand characteristics are not of

the same consequence, such as with archival research and non-obtrusive measures. However, in those instances in which a plausible alternative explanation for the laboratory phenomenon is offered, it is the obligation of the researcher to devise new means to study the phenomenon free of those alternative interpretations.

## REFERENCES


Adair, J. G., & Schacter, B. S. (1972). To cooperate or to look good? Subjects' and experimenter's perceptions of each others intentions. Journal of Experimental Social Psychology, 8, 74-85.

Arkes, H. R., Sarville, P. D., Wortmann, R. L., & Harkness, A. R. (1981). Hindsight bias among physicians weighting the likelihood of diagnoses. Journal of Applied Psychology, 65, 252-254.

Asch, S. E. (1952). Social Psychology. Englewood Cliffs: Prentice-Hall, Inc.

Briggs, S. R., Cheek, J. M., & Buss, A. H. (1980). An analysis of the Self-Monitoring Scale. Journal of Personality and Social Psychology, 38, 679-686.

Campbell, J., & Tesser, A. (1982). Hindsight bias: Why persons thought they knew it all along. Unpublished manuscript, University of Georgia, Atlanta.

Campbell, J., & Tesser, A. (1983). Motivational interpretations of hindsight bias: An individual difference analysis. Journal of Personality, 51, 605-621.

Cook, R. D. (1979). Influential observations in linear regression. Journal of the American Statistical Association, 74, 169-174.

Fischhoff, B. (1975). Hindsight=foresight: The effect of outcome knowledge on judgement under uncertainty. Journal of Experimental Psychology: Human Perception and Performance, 1, 288-299.

Fischhoff, B. (1977). Perceived informativeness of facts. Journal of Experimental Psychology, 3, 349-358.

Fischhoff, B. (1982). For those condemned to study the past: Heuristics and biases in hindsight. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgement under uncertainty: Heuristics and biases. (pp. 335-351). New York:Cambridge University Press.

Fischhoff, B., & Beyth, R. (1975). "I knew it would happen"- Remembered probabilities of once-future things. Organizational Behavior and Human Performance, 13, 1-16.

Gilovich, T. (1981). See the past in the present: The effect of associations to familiar events on judgments and decisions. Journal of Personality and Social Psychology, 40, 797-808.

Goranson, R. D., (1985). The hindsight effect in problem solving. Unpublished manuscript, York University, Downsview.

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. Journal of Verbal Learning and Verbal Behavior, 16, 107-112.

Hasher, L., Attig, M. S., & Abla, J. W. (1979). I knew it all along: Or, did I? Journal of Verbal Learning and Verbal Behavior, 20, 86-91.

Helmreich, R., & Stapp, J. (1974). Short forms of the Texas Social Behavior Inventory (TSBI): An objective measure of self-esteem. Bulletin of the Psychonomic Society, 4, 473-475.

Himmelfarb, S. (1975). What do you do when the control group doesn't fit into the factorial design? Psychological Bulletin, 82, 363-68.

Janoff-Bulman, R., Timko, C., & Carli, L. L. (1985). Cognitive biases in blaming the victim. Journal of Experimental Social Psychology, 21, 161-175.

Mitchell, T. R., & Kalb, L. S. (1981). Effect of outcome knowledge and outcome valence on supervisors' evaluations. Journal of Applied Psychology, 66, 604-612.

Moore, H. T. (1921). The comparative influence of majority and expert opinion. American Journal of Psychology, 32, 16-20.

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

Orne, M. T. (1970). Hypnosis, motivation, and the ecological validity of the psychological experiment. in W. J. Arnold & M. M. Page (Eds.), Nebraska Symposium on Motivation. Lincoln: University of Nebraska.

Page, M. M. (1973). On detecting demand awareness by post-experimental questionnaire. Journal of Social Psychology, 91, 305-323.

Pennington, D. C. (1981). The British firemen's strike of 1977/78. An investigation of judgements in foresight and hindsight. British Journal of Social Psychology, 20, 89-96.

Pennington, D. C., Rutter, D. R., McKenna, K., & Morley, I. E. (1980). Estimating the outcome of a pregnancy test: Women's judgements in foresight and hindsight. British Journal of Social and Clinical Psychology, 19, 317-324.

Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. Journal of Clinical Psychology, 38, 119-125.

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. Journal of Personality and Social Psychology, 32, 880-892.

Ross, L., Lepper, M. R., Strack, F., & Steinmetz, J. (1977). Social explanation and social expectation: Effects of real and hypothetical explanations on subjective likelihood. Journal of Personality and Social Psychology, 35, 817-829.

Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. Journal of Experimental Psychology, 3, 544-551.

Tabachnick, B.G., & Fidell, L.S. (1983). Using multivariate statistics. New York: Harper & Row Publishers.

Wood, G. (1978) The knew-it-all-along effect. Journal of Experimental Psychology: Human Perception and Performance, 4 345-353.

## Appendix A

## FISCHHOFF'S DEBIASING INSTRUCTIONS

On the following pages you will find a number of additional items which we intend to use in a subsequent study, identical to the one in which you have just participated. Although the correct answers to these items are indicated by a circle, we would like you to respond to them as you believe you would have responded had you not been told the answer. Your responses will enable us to evaluate the perceived difficulty of these items.

On previous occasions in which we have given people this task, we have found that they exaggerate how much they have known without being told the answer. You might call this an I-knew-it-all-along effect.

Consider, for example, the following question. Adaptive radiation refers to (a) evolutionary changes in animal life toward increased specialization or (b) the movement of of animals to a more suitable environment for survival. A group of people who were told that the correct answer was a believed that they would have assigned a probability of about .60 to a. A group of people who were not told the answer believed that the item was a toss-up. They assigned a probability of .50 to a. Another group of people who were told that the correct answer was a believed that they would have assigned a probability of .40 to b, the incorrect answer. Again, people who were not told the answer assigned a probability of .50 to b. As you can see, people who were told the answer to an item assigned a higher probability to the correct answer or a lower probability to the incorrect answer than they might have if they had not been told the answer.

In completing the present questionnaire, please do everything you can to avoid this bias. One reason why it happens is that people who are told the correct answer find it hard to imagine how they ever could have believed in the incorrect one. In answering, make certain that you haven't forgotten any reasons that you might have though of in favor of the wrong answer-had you not been told that it was wrong. In addition to figuring out how the correct answer fits in with whatever else you know about each topic, devote some attention to trying to see how the incorrect answer might also have fit in.

At the other extreme, however, be careful not to overcorrect and sell yourself short by underestimating how much you would have known without the answer

Appendix B

EXPERIMENTAL BOOKLETS

Sex: M___ F___

Native English ___ Other ___

Instructions

The purpose of this experiment is to examine the way individuals use the general knowledge that they have acquired in arriving at answers to new questions. To achieve this goal, we ask you to respond to the following general knowledge items. The items are ones for which we expect most people will not have exact knowledge of the correct answers. However, they are questions for which most of you can make more or less educated guesses based on your store of very general knowledge.

For each item indicate whether you believe the statement is true or false and the degree of certainty you feel in your answer. This is to be done by marking a 9-point line scale anchored by "certainly false" and "certainly true". The midpoint indicates "don't know at all". For example:

Albany is the capital of New York State.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

This is true, and if you knew this to be so, you would have circled a value between +1 and +4 depending on your certainly that the statement was correct. If you did not know the statement was true, you would have circled the value of 0. If you had believed that the statement was false, you would have circled a value between -1 and -4, depending on your certainty that the statement was incorrect.

When you have completed the forty items, please stop, put the booklet face down on your desk, and wait for further instructions.

1. In chess the queen is the only piece that can jump other pieces.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

2. Australia is approximately equal in area to the continental United States.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

3. The People's Republic of China was founded in 1947.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

4. The adult human body has some 200 miles of blood vessels.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

5. Dr. David Livingston was opposed to the slave trade in Africa.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

6. The movie, "The Godfather", grossed nearly 50 million dollars in its first year.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

7. Cairo, Egypt has a larger population than Chicago, Illinois.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

8. The second closest star to our solar system is known as Barnard's star.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

9. In the world, there are more Roman Catholics than there are Moslems.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

10. The capybara is the largest of the marsupials.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

11. Irving Berlin wrote the song called "Oh how I hate to get up in the morning."

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

12. Groups of ants can cross streams by linking bodies and swimming.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

13. Lithium is the lightest of all metals.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

14. The New York Rangers last won the Stanley Cup in 1960.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

15. Pablo Picasso was the most prolific of all known painters.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

16. The largest religious building in the world is Ang Kor Wat in Cambodia.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

17. Diving did not become part of the Olympic program until 1946.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

18. Divorce is found only in technologically advanced societies.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

19. Tin is the traditional gift associated with the first wedding anniversary.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

20. The Aurora Borealis appears in the southern hemisphere.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

21. Between 1917 and 1919 a world influenza epidemic killed nearly a million people.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

22. The origin of the term Hinduism is the Indian word for law.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

23. The Panama canal is about 70 miles long.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

24. Mexico has the largest population of any Latin American country.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

25. Brass is an alloy of tin and copper.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

26. Michaelangelo did a famous painting called "Still life with apples."

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

27. The Bible has been banned more frequently through history than any other book.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

28. At temperatures close to absolute zero, the magnetic properties of many substances undergo change.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

29. The green part of a sprouting potato is poisonous.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

30. World temperance day is celebrated each year on the first Sunday in October.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

31. The movie "Casablanca" won the Academy Award for best picture in 1943.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

32. In Malaysia, if a man goes to jail for being drunk, his wife goes too.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

33. Fortune cookies originally contained Bible passages.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

34. It takes twice as much fuel to move a ton of freight by railroad as it does by truck.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

35. Floyd Patterson was the youngest man to win the world heavy weight boxing championship.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

36. Rocky Marciano was the only heavyweight champion ever to retire without losing a professional fight.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

37. The equatorial radius of Mars is approximately equal to that of earth.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

38. The Scopes trial, often referred to as the Monkey Trial, took place in 1934.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

39. Denmark has the highest suicide rate of any country in the world.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

40. The busiest airport in the world is Orly Airport in Paris.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

STOP

Please answer the following items on the IBM sheet that is provided. For each item, indicate how characteristic each of the statements are about you on a scale from 1 to 5.

1 = Not at all characteristic of me

2 = Not very characteristic of me

3 = Slightly characteristic of me

4 = Fairly characteristic of me

5 = Very characteristic of me

1. Before voting I thoroughly investigate the qualifications of all the candidates.

2. I never hesitate to go out of my way to help someone in trouble.

3. It is sometimes hard for me to go on with my work if I am not encouraged.

4. I have never intensely disliked anyone.

5. On occasion I have had doubts about my ability to succeed in life.

6. I sometimes feel resentful when I don't get my way.

7. I am always careful about my manner of dress.

8. My table manners at home are as good as when I eat out in a restaurant.

9. If I could get into a movie without paying and be sure I was not seen I would probably do it.

10. On a few occasions, I have given up doing something because I thought too little of my ability.

11. I like to gossip at times.

12. There have been times when I felt like rebelling against people in authority even though I knew they were right.

13. No matter who I'm talking to, I'm always a good listener.

14. I can remember "playing sick" to get out of something.

15. There have been occasions when I took advantage of someone.

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

17. I always try to practice what I preach.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

33. I have never deliberately said something that hurt someone's feelings.

34. I am not likely to speak to people until they speak to me.

35. I would describe myself as self-confident.

36. I feel confident of my appearance.

37. I am a good mixer.

38. When in a group of people, I have trouble thinking of the right things to say.

39. When in a group of people, I usually do what the others want rather than make suggestions.

40. When I am in disagreement with other people, my opinion usually prevails.

41. I would describe myself as one who attempts to master situations.

42. Other people look up to me.

43. I enjoy social gatherings just to be with people.

44. I make a point of looking other people in the eye.

45. I cannot seem to get others to notice me.

46. I would rather not have very much responsibility for other people.

47. I feel comfortable being approached by someone in a position of authority.

48. I would describe myself as indecisive.

49. I have no doubts about my social competence.

50. I would describe myself as socially unskilled.

51. I would be willing to describe myself as a pretty "strong" personality.

52. When I work on a committee I like to take charge of things.

53. I usually expect to succeed in the things I do.

54. I feel comfortable approaching someone in a position of authority over me.

55. I enjoy being around other people, and seek out social encounters frequently.

56. I feel confident of my social behavior.

57. I feel I can confidently approach and deal with anyone I meet.

58. I would describe myself as happy.

59. I enjoy being in front of large audiences.

60. When I meet a stranger, I often think that he is better than I am.

61. It is hard for me to start a conversation with strangers.

62. People seem naturally to turn to me when decisions have to be made.

63. I feel secure in social situations.

64. I like to exert my influence over other people.

65. My behavior is usually an expression of my true inner feelings, attitudes, and beliefs.

66. At parties and social gatherings, I do not attempt to do or say things that others will like.

67. I can make impromptu speeches on topics about which I have almost no information.

68. I guess I put on a show to impress or entertain people.

69. When I am uncertain how to act in social sitations, I look to the behavior of others for cues.

70. I would probably make a good actor.

71. In a group of people I am rarely the center of attention.

72. In different situations and with different people, I often act like very different persons.

73. I am not particularly good at making other people like me.

74. Even if I am not enjoying myself, I often pretend to be having a good time.

75. I'm not always the person I appear to be.

76. I would not change my opinions (or the way I do things) in order to please someone else or win their favor.

77. I have considered being an entertainer.

78. In order to get along and be liked, I tend to be what people expect me to be rather than anything else.

79. I have never been good at games like charades or improvisational acting.

80. I have trouble changing my behavior to suit different people and different situations.

81. At a party I let others keep the jokes and stories going.

82. I feel a bit awkward in company and do not show up quite as well as I should.

83. I can look anyone in the eye and tell a lie with a straight face (if for a right end).

On the following pages you will find a number of additional statements. Although the correct answers to these items are indicated, we would like you to respond to them as you believe you would have responded had you not been told the answer.

On previous occasions in which we have given people this task, we have found that they underestimate (overestimate) how much they would have known without being told the answer. You might call this an 'I-never-knew-that' ('I-knew-it-all-along') effect.

Consider, for example, the following question. "Adaptive radiation refers to evolutionary changes in animal life toward increased specialization." A group of people who were told that this was a true statement assigned a confidence rating of -1 (+1) on the nine point scale. A group of people who were not told this believed that the item was a toss-up. They assigned a confidence rating of 0 to the item. Another group of people who were told that this was a false statement assigned a confidence rating of +1 (-1) to the item. Again, people who were not told the answer assigned a confidence rating of 0 to the item. As you can see, people who were told the answer to an item assigned a lower (higher) rating to the true statement or a higher (lower) rating to the false statement than they might have if they had not been told the answer. In completing the present questionnaire, please do everything you can to avoid this bias.

( One reason why it happens is that people who are told the answer find it hard to imagine how they ever could have believed that the item was true. In answering, make certain that you have not forgotten any reasons that you might have thought of in favor of rating the item the other way, had you not been told the answer was true. In addition to figuring out how the correct answer fits in with whatever else you know about each topic, devote some attention to trying to see how the answer might also have been the other way.

At the other extreme, however, be careful not to undercorrect and sell yourself short by overestimating how much you would have known without the answer. )

## Section 2

FOR THE FOLLOWING 20 ITEMS, RECALL AS ACCURATELY AS POSSIBLE THE RESPONSE YOU GAVE EARLIER TO EACH STATEMENT. THAT IS, MARK THE SCALE FOR EACH ITEM AS YOU BELIEVE YOU MARKED IT THE FIRST TIME YOU RESPONDED.

1. Michaelangelo did a famous painting called "Still life with apples". (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

2. Brass is an alloy of tin and copper. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

3. Lithium is the lightest of all metals. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

4. Mexico has the largest population of any Latin American country. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

5. Between 1917 and 1919 a world influenza epidemic killed nearly a million people. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

6. Irving Berlin wrote the song called "Oh how I hate to get up in the morning". (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

7. The adult human body has some 200 miles of blood vessels. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

8. Groups of ants can cross streams by linking bodies and swimming. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

9. The Bible has been banned more frequently through history than any other book. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

10. Dr. David Livingston was opposed to the slave trade in Africa. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

11. Tin is the traditional gift associated with the first wedding anniversary. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

12. It takes twice as much fuel to move a ton of freight by railroad as it does by truck. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

13. Fortune cookies originally contained Bible passages. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

14. The green part of a sprouting potato is poisonous. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

15. The largest religious building in the world is Ang Kor Wat in Cambodia. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

16. The Aurora Borealis appears in the southern hemisphere. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

17. The People's Republic of CHina was founded in 1947. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

18. Cairo, Egypt has a larger population than Chicago, Illinois. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

19. The capybara is the largest of the marsupials. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

20. The movie, "The Godfather", grossed nearly 50 million dollars in its first year. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

## Section 3

FOR THE FOLLOWING 20 ITEMS, TRY TO ESTIMATE AS ACCURATELY AS YOU CAN THE ANSWER YOU BELIEVE YOU WOULD HAVE GIVEN IF WE HAD NOT INDICATED THE CORRECT ANSWER. THAT IS, MARK THE SCALE FOR EACH ITEM AS YOU BELIEVE YOU WOULD HAVE MARKED IT HAD YOU NOT BEEN TOLD THE CORRECT ANSWER.

1. London's Bobbies were named after British statesman, Sir Robert Peale. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

2. The first modern Olympic games were held in Rome, Italy. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

3. Rich women are more likely to get married than poor women. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

4. Canada is the world's leading producer of silver. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

5. The thigh bone is the longest bone in the human body. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

6. The largest museum in the world is the Louvre in Paris. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

7. A rock more than 10 inches in diameter is called a boulder. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

8. The Atlantic Ocean is deeper on the average than the Pacific. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

9. There are no doorknobs in Iceland. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

10. Females commit suicide more frequently than do males. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

11. Ultra-violet radiation is capable of killing bacteria. (True) (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

12. The total population of Greenland is about 50,000. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

13. Krypton is found naturally in crystalline form. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

14. The maximum life span of the chicken is about five years. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

15. The inner core of the earth is composed largely of liquid metals. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

16. In 1973, tropical storm Agnes caused 3.5 billion dollars in damage. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

17. The largest dam in the world is in Pakistan. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

18. The Indian Ocean is the smallest ocean on the earth. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

19. In the movie, "The Petrified Forest" Humphrey Bogart plays a criminal. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

20. The minimum voting age in the USSR is 18. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

In this section, I am interested in determining your understanding of the 'I-never-knew-that' effect as it was just presented. You will be asked to rate a few more statements. Some of these statements you have seen previously, some you have not. What I want you to do now is to rate these statements as you have been, except this time I want you to purposely rate some of the statements as if you were <u>under</u> the influence of the 'I-never-knew-that' effect, and some as if you were <u>not</u> under the influence of the 'I-never-knew-that' effect. Read the directions that precede each of the following statements carefully, and rate each statement according to those directions as best you can.

## Section 4

RATE THESE STATEMENTS AS A SUBJECT WOULD WHO WAS UNDER THE INFLUENCE OF THE 'I-NEVER-KNEW-THAT' EFFECT.

1. The busiest airport in the world is Orly Airport in Paris. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

2. Australia is approximately equal in area to the continental United States. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

3. The Scopes trial, often referred to as the Monkey Trial, took place in 1934. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

4. In the world, there are more Roman Catholics than there are Moslems. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

5. The equatorial radius of Mars is approximately equal to that of earth. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

6. The origin of the term Hinduism is the Indian word for law. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

7. At temperatures close to absolute zero, the magnetic properties of many substances undergo change. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

8. In chess the queen is the only piece that can jump other pieces. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

9. The movie "Casablanca" won the Academy Award for best picture in 1943. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

10. The New York Rangers last won the Stanley Cup in 1960. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

11. Coffee got its name from the Ethiopian province of Kaffa. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

12. The maximum duration of a solar eclipse is 17 minutes. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

13. Ernest Hemingway received a Pulitzer Prize for The Old Man and the Sea. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

14. The stegosaurus had a brain which weighed up to 70 pounds. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

15. Pele has scored more goals than any other professional soccer player. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

16. Antioch was a popular city in ancient Iraq. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

17. Shintoism is the native religion of Korea. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

18. For unknown reasons alcoholics almost never go bald. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

19. Fauvist art can be distinguished by its extremely muted colors. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

20. Blowing up a balloon is a good test for assessing the vital capacity of the lungs. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

## Section 5

RATE THESE STATEMENTS AS A SUBJECT WOULD WHO WAS NOT UNDER THE INFLUENCE OF THE 'I-NEVER-KNEW-THAT' EFFECT.

1. In Malaysia, if a man goes to jail for being drunk, his wife goes too. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

2. Diving did not become part of the Olympic program until 1946. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

3. Rocky Marciano was the only heavyweight champion ever to retire without losing a professional fight. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

4. The second closest star to our solar system is known as Barnard's star.

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

5. World temperance day is celebrated each year on the first Sunday in October. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

6. Denmark has the highest suicide rate of any country in the world. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

7. The Panama canal is about 70 miles long. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

8. Divorce is found only in technologically advanced societies. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

9. Pablo Picasso was the most prolific of all known painters. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

10. Floyd Patterson was the youngest man to win the world heavy weight boxing championship. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

11. Babe Ruth has the highest lifetime batting average in professional baseball. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

12. Precipitation is very light in much of the Arctic region. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

13. Marlon Brando has won the Academy Award for best actor three times. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

14. The Salk polio vaccine was discovered in the 1940's. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

15. Laughing sickness can be a fatal disease. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

16. High does of Vitamin A can produce blurring of vision and dizziness. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

17. The Queen Elizabeth cup is awarded for the best show horse at the Royal International Horse Show. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

18. Earth is the only planet in this solar system with one moon. (True)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

19. In most eastern cultures, men adopt the women's family following marriage. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

20. The largest denomination of currency in circulation is the 5000 dollar bill. (False)

[Certainly False] -4 -3 -2 -1 0 +1 +2 +3 +4 [Certainly True]

## Appendix C

## POST-EXPERIMENTAL QUESTIONNAIRE

The results of an experiment are more meaningful to us if we know what your ideas, thoughts and understandings of the experiment were. Please answer each of the questions on the following pages frankly and honestly. Please answer them in their numbered order and do not go on to the next question until you have given an answer to the previous question. Do not go back to a question once you have started on the next one. Remember, we want you to answer the questions as accurately as you can.

1. What did you think the experimenter expected to find in this study?

( Aware 33  Unaware 44  Missing 3 )

2. What do you think the reason was for twice rating the same items?

( Aware 50  Unaware 37  Missing 3 )

3. For those statements that you rated twice, how did you think you were expected to rate the second set of items as compared with the first set?

( Aware 35  Unaware 41  Missing 4 )

4. For those statements labelled as false, how did you think you were expected to rate the second set of items as compared with the first set?

[Closer to the extreme] 1 2 3 4 5 6 7 8 9 [Closer to the middle]

( $\underline{M}$ = 4.84 )

5. For those statements labelled as true, how did you think you were expected to rate the second set of items as compared with the first set?

[Closer to the extreme] 1 2 3 4 5 6 7 8 9 [Closer to the middle]

( $\underline{M}$ = 5.01 )

6. Did you find it more difficult to rate the statements on which you were told the answers than to rate those statements that you were not told the answers for?

[Not at all] 1 2 3 4 5 6 7 8 9 [Very much so]

( $\underline{M}$ = 5.54 )

7. What effect was being provided with the answer supposed to have on your ratings?

( Aware 55  Unaware 20  Missing 5 )

a) Did it have that effect?

( Yes 59  No 15  Missing 6 )

8. Did you do anything to try to avoid the bias of having been provided with the answers to some of the items? If so, what?

( Yes 57  No 19  Missing 4 )

9. Did you feel that the answers to some of the statements were not correct?

[Not at all] 1 2 3 4 5 6 7 8 9 [Very much so]

( $\underline{M}$ = 4.88 )

10. Did you think that rerating the same set of items twice was a test of your honesty?

[Not at all] 1 2 3 4 5 6 7 8 9 [Very much so]

( $\underline{M}$ = 5.6 )

11. Do you feel these items were designed to measure your intelligence?

[Not at all] 1 2 3 4 5 6 7 8 9 [Very much so]

( $\underline{M}$ = 3.13 )

Note: a. For items 1, 2, 3 and 7a, a response was classified as aware if judged to be congruent with awareness as to the purpose of the outcome feedback.

b. For items 4, 5, 6, 9, 10 and 11, the mean values are reported.

## Appendix D

## DEBRIEFING

The study you have participated in concerns the hindsight bias. The hindsight bias is the tendency to misjudge the extent one would have known the solution to a problem after the solution is revealed. For instance, after being told that the correct answer to a test item is 'C', and having mistakenly selected 'B', there is a tendency to believe one should have known to have picked the answer 'C'. The present study explores three possible sources of the bias. The bias may be the product of incorporating outcome information into the knowledge previously possessed about the topic. The bias may be the product of motivational factors, such as a desire to appear to others to have known the correct answer to the problem. The bias may be the product of recognizing that the outcome information provided is provided for a purpose, and altering the ratings given accordingly. The bias may be the combination of all of these factors.

After completing the personality measures, but prior to rating items with the outcome information provided, some of you were told that our interest was in exploring what was

termed the 'Knew-it all-along' effect, the 'Never-knew-that' effect, or you were not provided with a description of the phenomenon. Furthermore, some of you were provided with a more complete description of the bias than were others. Again, our interest was in determining the extent to which these instructions may influence the bias. Our interest was not in whether you knew the correct answers to the statements. The statements were selected such that it would be unlikely that you would know for certain the correct solution to many of them.

# Appendix E

## SUMMARY OF ANOVAS AND CELL MEANS, DEBIASING PHASE

Table E-1

ANOVA Summary Table, Debiasing Phase, Dependent Variable True-Hypothetical

| Variable(s) | F | p |
|---|---|---|
| Set (S) | 3.19 | .0749 |
| Direction (D) | .09 | .7681 |
| Order (O) | 1.34 | .2479 |
| Length (L) | .50 | .4821 |
| SxD | .52 | .4728 |
| SxL | .03 | .8619 |
| SxO | .63 | .4275 |
| DxL | .05 | .8265 |
| DxO | .61 | .4363 |
| LxO | .06 | .8030 |
| SxDxL | .39 | .5304 |
| SxDxO | .78 | .3767 |
| SxLxO | .19 | .6665 |
| LxOxD | .34 | .5606 |
| SxDxLxO | 2.70 | .1016 |

Note: degrees of freedom for each test were 1 and 304.

Table E-2

ANOVA Summary Table, Debiasing Phase, Dependent Variable False-Hypothetical

| Variable(s) | F | p |
|---|---|---|
| Set (S) | .98 | .3218 |
| Direction (D) | .91 | .3421 |
| Order (O) | 1.30 | .2551 |
| Length (L) | .22 | .6430 |
| SxD | 3.53 | .0611 |
| SxL | 1.64 | .2009 |
| SxO | .00 | .9687 |
| DxL | 1.52 | .2179 |
| DxO | .45 | .5043 |
| LxO | 1.23 | .2684 |
| SxDxL | .39 | .5349 |
| SxDxO | 1.23 | .2684 |
| SxLxO | .37 | .5453 |
| LxOxD | .79 | .3749 |
| SxDxLxO | .46 | .6333 |

Note: degrees of freedom for each test were 1 and 304.

Table E-3

ANOVA Summary Table, Debiasing Phase,
Dependent Variable True-Memory

| Variable(s) | F | p |
|---|---|---|
| Set (S) | .73 | .3928 |
| Direction (D) | 1.26 | .2626 |
| Order (O) | .32 | .5713 |
| Length (L) | .00 | .9911 |
| SxD | 1.75 | .1870 |
| SxL | .00 | .9911 |
| SxO | 1.11 | .2919 |
| DxL | .02 | .8852 |
| DxO | .04 | .8329 |
| LxO | .05 | .8156 |
| SxDxL | .43 | .5125 |
| SxDxO | .43 | .5125 |
| SxLxO | 1.02 | .3127 |
| LxOxD | .02 | .8852 |
| SxDxLxO | 1.58 | .2102 |

Note: degrees of freedom for each test were 1 and 304.

Table E-4

ANOVA Summary Table, Debiasing Phase,
Dependent Variable False-Memory

| Variable(s) | F | p |
|---|---|---|
| Set (S) | 9.94 | .0018 |
| Direction (D) | .47 | .4932 |
| Order (O) | 2.72 | .0099 |
| Length (L) | .07 | .7953 |
| SxD | 2.15 | .1440 |
| SxL | .20 | .6566 |
| SxO | .03 | .8676 |
| DxL | .03 | .8679 |
| DxO | .61 | .4367 |
| LxO | .18 | .6700 |
| SxDxL | .21 | .6413 |
| SxDxO | .93 | .3357 |
| SxLxO | 1.12 | .2914 |
| LxOxD | 2.09 | .1491 |
| SxDxLxO | .20 | .6566 |

Note: degrees of freedom for each test were 1 and 304.

Table E-5

Interaction Cell Means for the Four Measures of the Hindsight Bias, Debiasing Phase

| Variables | Hypothetical | | Memory | |
|---|---|---|---|---|
| | True | False | True | False |
| Set x Direction | | | | |
| A Know | .25 | -.86 | -.00 | -.14 |
| A Never | .34 | -.64 | .14 | -.09 |
| B Know | .46 | -.63 | .02 | -.26 |
| B Never | .42 | -.71 | .01 | -.40 |
| Set x Length | | | | |
| A Complete | .27 | -.78 | .07 | -.09 |
| A Shortened | .32 | -.71 | .07 | -.14 |
| B Complete | .46 | -.72 | -.03 | -.38 |
| B Shortened | .48 | -.74 | .02 | -.32 |
| Set x Order | | | | |
| A H/M | .37 | -.79 | .08 | -.18 |
| A M/H | .21 | -.71 | .05 | -.05 |
| B H/M | .46 | -.72 | -.03 | -.38 |
| B M/H | .43 | -.62 | .06 | -.28 |
| Direction x Length | | | | |
| Knew Complete | .34 | -.78 | .01 | -.20 |
| Knew Shortened | .38 | -.72 | .02 | -.20 |
| Never Complete | .34 | -.60 | .08 | -.23 |

| | | | | |
|---|---|---|---|---|
| Never Shortened | .42 | -.74 | .07 | -.26 |
| Direction x Order | | | | |
| Knew H/M | .44 | -.82 | -.01 | -.23 |
| Knew M/H | .28 | -.68 | .03 | -.17 |
| Never H/M | .40 | -.69 | .06 | -.33 |
| Never M/H | .36 | -.65 | .08 | -.16 |
| Length x Order | | | | |
| Complete H/M | .40 | -.78 | .03 | -.25 |
| Complete M/H | .28 | -.60 | .05 | -.17 |
| Shortened H/M | .43 | -.73 | .02 | -.30 |
| Shortened M/H | .36 | -.73 | .07 | -.16 |
| SxDxL | | | | |
| A Knew Complete | .21 | -.97 | -.02 | -.14 |
| A Knew Shortened | .29 | -.76 | .02 | -.14 |
| A Never Complete | .33 | -.60 | .16 | -.04 |
| A Never Shortened | .34 | -.68 | .11 | -.13 |
| B Knew Complete | .46 | -.59 | .04 | -.25 |
| B Knew Shortened | .46 | -.68 | .01 | -.26 |
| B Never Complete | .35 | -.61 | -.00 | -.42 |
| B Never Shortened | .50 | -.80 | .03 | -.38 |
| SxDxO | | | | |
| A Knew H/M | .33 | -.88 | .02 | -.21 |
| A Knew M/H | .18 | -.84 | -.03 | -.07 |
| A Never H/M | .42 | -.70 | .14 | -.14 |
| A Never M/H | .25 | -.58 | .13 | -.03 |
| B Knew H/M | .54 | -.75 | -.04 | -.25 |
| B Knew M/H | .38 | -.52 | .09 | -.26 |

| | | | | |
|---|---|---|---|---|
| B Never H/M | .37 | -.68 | -.01 | -.51 |
| B Never M/H | .48 | -.73 | .03 | -.29 |
| SxLxO | | | | |
| A Complete H/M | .38 | -.85 | .11 | -.17 |
| A Complete M/H | .16 | -.72 | .02 | -.01 |
| A Shortened H/M | .37 | -.74 | .05 | -.18 |
| A Shortened M/H | .26 | -.69 | .09 | -.10 |
| B Complete H/M | .41 | -.72 | -.05 | -.33 |
| B Complete M/H | .40 | -.49 | .08 | -.33 |
| B Shortened H/M | .50 | -.72 | -.01 | -.42 |
| B Shortened M/H | .46 | -.76 | .04 | -.22 |
| DxLxO | | | | |
| Knew Comp H/M | .40 | -.93 | -.00 | -.16 |
| Knew Comp M/H | .27 | -.63 | .02 | -.23 |
| Knew Short H/M | .47 | -.71 | -.01 | -.29 |
| Knew Short M/H | .28 | -.72 | .05 | -.11 |
| Never Comp H/M | .39 | -.63 | .07 | -.35 |
| Never Comp M/H | .29 | -.58 | .08 | -.11 |
| Never Short H/M | .40 | -.75 | .06 | -.31 |
| Never Short M/H | .44 | -.73 | .08 | -.21 |
| SxDxLxO | | | | |
| A Knew Comp H/M | .23 | -1.04 | .00 | -.16 |
| A Knew Comp M/H | .20 | -.90 | -.06 | -.11 |
| A Knew Short H/M | .43 | -.74 | .04 | -.26 |
| A Knew Short M/H | .15 | -.77 | .00 | -.03 |
| A Never Comp H/M | .54 | -.66 | .22 | -.18 |
| A Never Comp M/H | .13 | -.54 | .09 | .09 |

| | | | | |
|---|---|---|---|---|
| A Never Short H/M | .30 | -.74 | .05 | -.10 |
| A Never Short M/H | .38 | -.62 | .18 | -.16 |
| B Knew Comp H/M | .58 | -.82 | -.01 | -.16 |
| B Knew Comp M/H | .34 | -.35 | .09 | -.34 |
| B Knew Short H/M | .51 | -.68 | -.07 | -.33 |
| B Knew Short M/H | .41 | -.68 | .10 | -.18 |
| B Never Comp H/M | .25 | -.61 | .08 | -.51 |
| B Never Comp M/H | .46 | -.62 | .08 | -.32 |
| B Never Short H/M | .49 | -.76 | .07 | -.51 |
| B Never Short M/H | .50 | -.84 | -.01 | -.26 |

# Appendix F

## SUMMARY OF ANOVAS AND CELL MEANS, INTENTIONAL-BIAS PHASE

Table F-1

ANOVA Summary Table for the Intentional-Bias Phase, Dependent Variable True-Hypothetical

| Variable(s) | F | p |
|---|---|---|
| Set (S) | 7.39 | .0069 |
| Direction (D) | 44.06 | .0001 |
| Order (O) | .11 | .7387 |
| Length (L) | .80 | .3716 |
| Understanding (U) | 8.15 | .0016 |
| SxD | .48 | .4879 |
| SxL | .14 | .7087 |
| SxO | .35 | .5572 |
| DxL | .46 | .4963 |
| DxO | .00 | .9894 |
| LxO | .58 | .4471 |
| LxS | 3.15 | .0769 |
| UxD | 90.31 | .0001 |
| UxL | .76 | .3837 |
| UxO | .36 | .5511 |
| SxDxL | .16 | .6889 |
| SxDxO | .48 | .4879 |
| SxLxO | .35 | .5572 |
| LxOxD | .24 | .6216 |
| SxDxU | .18 | .6686 |

| | | |
|---|---|---|
| SxLxU | 6.77 | .0097 |
| SxOxU | 1.65 | .1196 |
| LxOxU | 2.39 | .1232 |
| DxLxU | 1.86 | .1742 |
| DxOxU | 1.32 | .2520 |
| SxDxLxO | 2.74 | .0987 |
| SxDxLxU | .41 | .5209 |
| SxDxOxU | .34 | .5613 |
| SxLxOxU | 3.60 | .0587 |
| DxLxOxU | .47 | .4916 |
| SxDxLxOxU | 3.26 | .0719 |

Note: degrees of freedom for each test were 1 and 304.

Table F-2

ANOVA Summary Table for the Intentional-Bias Phase, Dependent Variable False-Hypothetical

| Variable(s) | F | p |
|---|---|---|
| Set (S) | .25 | .6143 |
| Direction (D) | 15.04 | .0001 |
| Order (O) | 2.56 | .1106 |
| Length (L) | 3.11 | .0789 |
| Understanding (U) | .31 | .5768 |
| SxD | 5.01 | .0260 |
| SxL | 5.63 | .0183 |
| SxO | .37 | .5422 |
| DxL | .31 | .5787 |
| DxO | .00 | .9892 |
| LxO | .08 | .7760 |
| UxS | .42 | .5185 |
| UxD | 51.73 | .0001 |
| UxL | .00 | .9890 |
| UxO | .21 | .6503 |
| SxDxL | 1.12 | .2911 |
| SxDxO | 2.47 | .1168 |
| SxLxO | .80 | .3715 |
| LxOxD | .36 | .5512 |
| SxDxU | 5.09 | .0248 |

| | | |
|---|---|---|
| SxLxU | .02 | .8907 |
| SxOxU | .63 | .7209 |
| LxOxU | 1.97 | .1617 |
| DxLxU | .08 | .7835 |
| DxOxU | .01 | .9343 |
| SxDxLxO | .28 | .5973 |
| SxDxLxU | .00 | .9671 |
| SxDxOxU | 1.30 | .2546 |
| SxLxOxU | .66 | .4178 |
| DxLxOxU | .29 | .5921 |
| SxDxLxOxU | .33 | .5640 |

Note: degrees of freedom for each test were 1 and 304.

Table F-3

ANOVA Summary Table for the Intentional-Bias Phase, Dependent Variable True-Memory

| Variable(s) | F | p |
|---|---|---|
| Set (S) | 4.83 | .0287 |
| Direction (D) | 72.80 | .0001 |
| Order (O) | 2.11 | .1477 |
| Length (L) | 2.83 | .0936 |
| Understanding (U) | 46.74 | .0001 |
| SxD | .52 | .4727 |
| SxL | 2.23 | .1367 |
| SxO | 13.91 | .0002 |
| DxL | 2.27 | .1332 |
| DxO | 2.56 | .1105 |
| LxO | 2.78 | .0963 |
| UxS | 13.91 | .0002 |
| UxD | 67.43 | .0001 |
| UxL | .04 | .8427 |
| UxO | 28.59 | .0001 |
| SxDxL | 3.35 | .0680 |
| SxDxO | .50 | .4811 |
| SxLxO | 2.19 | .1403 |
| LxOxD | .46 | .4981 |
| SxDxU | .48 | .4876 |

| | | |
|---|---|---|
| SxLxU | .12 | .7338 |
| SxOxU | 7.19 | .0078 |
| LxOxU | 5.75 | .0171 |
| DxLxU | .52 | .4700 |
| DxOxU | .03 | .8650 |
| SxDxLxO | .05 | .8283 |
| SxDxLxU | .65 | .4195 |
| SxDxOxU | .29 | .5903 |
| SxLxOxU | .07 | .7877 |
| DxLxOxU | .99 | .3216 |
| SxDxLxOxU | .93 | .3356 |

Note: degrees of freedom for each test were 1 and 304.

Table F-4

ANOVA Summary Table for the Intentional-Bias Phase, Dependent Variable False-Memory

| Variable(s) | F | p |
|---|---|---|
| Set (S) | .97 | .3258 |
| Direction (D) | 52.57 | .0001 |
| Order (O) | 7.65 | .0060 |
| Length (L) | 3.73 | .0544 |
| Understanding (U) | 19.01 | .0001 |
| SxD | .10 | .7462 |
| SxL | .02 | .8911 |
| SxO | 12.52 | .0005 |
| DxL | 4.59 | .0329 |
| DxO | 1.65 | .2004 |
| LxO | 4.17 | .0419 |
| UxS | 10.19 | .0016 |
| UxD | 31.78 | .0001 |
| UxL | .98 | .3239 |
| UxO | 14.54 | .0002 |
| SxDxL | 1.95 | .1640 |
| SxDxO | .54 | .4629 |
| SxLxO | .03 | .8617 |
| LxOxD | .58 | .4479 |
| SxDxU | .06 | .8099 |

| | | |
|---|---|---|
| SxLxU | 1.24 | .2658 |
| SxOxU | 11.44 | .0008 |
| LxOxU | 3.47 | .0635 |
| DxLxU | .42 | .5187 |
| DxOxU | .47 | .4944 |
| SxDxLxO | 6.65 | .0104 |
| SxDxLxU | .42 | .5187 |
| SxDxOxU | 2.39 | .1233 |
| SxLxOxU | .03 | .8693 |
| DxLxOxU | .01 | .9395 |
| SxDxLxOxU | .79 | .3579 |

Note: degrees of freedom for each test were 4 and 301.

Table F-5

Cell Means for the Four Measures of the Hindsight Bias, Intentional-Bias Phase

| Variables | Hypothetical | | Memory | |
|---|---|---|---|---|
| | True | False | True | False |
| Set | | | | |
| A | .74 | -.74 | .44 | -.56 |
| B | .49 | -.76 | .23 | -.46 |
| Direction | | | | |
| Knew | .92 | -.93 | .73 | -.87 |
| Never | .30 | -.57 | -.06 | -.15 |
| Length | | | | |
| Complete | .57 | -.67 | .26 | -.41 |
| Shortened | .65 | -.83 | .41 | -.61 |
| Order | | | | |
| Hypo/Mem | .63 | -.69 | .40 | -.65 |
| Mem/Hypo | .60 | -.81 | .27 | -.37 |
| Understanding | | | | |
| Biased | .75 | -.73 | .65 | -.73 |
| Not Biased | .48 | -.78 | .02 | -.29 |
| SxD | | | | |
| A Know | 1.08 | -1.02 | .79 | -.94 |
| A Never | .40 | -.46 | .08 | -.18 |
| B Know | .76 | -.84 | .66 | -.81 |

| | | | | |
|---|---|---|---|---|
| B Never | .21 | -.69 | -.19 | -.11 |
| SxL | | | | |
| A Complete | .72 | -.77 | .43 | -.46 |
| A Shortened | .76 | -.71 | .44 | -.66 |
| B Complete | .43 | -.57 | .09 | -.37 |
| B Shortened | .54 | -.95 | .38 | -.55 |
| SxO | | | | |
| A H/M | .73 | -.65 | .33 | -.52 |
| A M/H | .75 | -.83 | .54 | -.60 |
| B H/M | .53 | -.73 | -.47 | -.78 |
| B M/H | .44 | -.80 | -.00 | -.15 |
| DxL | | | | |
| Knew Complete | .91 | -.88 | .58 | -.89 |
| Knew Shortened | .93 | -.98 | .87 | -.86 |
| Never Complete | .23 | -.47 | -.07 | .06 |
| Never Shortened | .38 | -.68 | -.05 | -.35 |
| DxO | | | | |
| Knew H/M | .94 | -.87 | .72 | -.95 |
| Knew M/H | .91 | -.99 | .73 | -.80 |
| Never H/M | .32 | -.51 | .08 | -.35 |
| Never M/H | .29 | -.64 | -.20 | .05 |
| LxO | | | | |
| Complete H/M | .62 | -.60 | .40 | -.45 |
| Complete M/H | .52 | -.75 | .11 | -.38 |
| Shortened H/M | .63 | -.79 | .40 | -.85 |
| Shortened M/H | .67 | -.88 | .42 | -.37 |
| SxU | | | | |

| | | | | |
|---|---|---|---|---|
| A Biased | .80 | -.75 | .60 | -.62 |
| A Not Biased | .68 | -.74 | .27 | -.50 |
| B Biased | .69 | -.71 | .70 | -.84 |
| B Not Biased | .28 | -.82 | -.23 | -.08 |
| DxU | | | | |
| Knew Biased | 1.44 | -1.23 | 1.40 | -1.37 |
| Knew Not Biased | .40 | -.63 | .05 | -.38 |
| Never Biased | .05 | -.22 | -.10 | -.09 |
| Never Not Biased | .56 | -.93 | -.01 | -.21 |
| LxU | | | | |
| Complete Biased | .67 | -.64 | -.56 | -.68 |
| Complete N.B. | .47 | -.70 | .05 | -.15 |
| Shortened Biased | .82 | -.81 | .74 | -.78 |
| Shortened N.B. | .49 | -.86 | .09 | -.44 |
| OxU | | | | |
| H/M Biased | .74 | -.64 | .48 | -.68 |
| H/M Not Biased | .52 | -.74 | .32 | -.62 |
| M/H Biased | .76 | -.81 | .82 | -.78 |
| M/H Not Biased | .44 | -.82 | -.28 | .03 |
| SxDxL | | | | |
| A Knew Complete | 1.11 | -1.13 | .80 | -1.01 |
| A Knew Shortened | 1.06 | -.92 | .79 | -.87 |
| A Never Complete | .32 | -.41 | .05 | .10 |
| A Never Shortened | .47 | -.51 | .10 | -.46 |
| B Knew Complete | .72 | -.62 | .36 | -.75 |
| B Knew Shortened | .81 | -1.05 | .96 | -.86 |
| B Never Complete | .13 | -.52 | -.18 | .01 |

| | | | | |
|---|---|---|---|---|
| B Never Shortened | .28 | -.86 | -.20 | -.24 |
| SxDxO | | | | |
| A Knew H/M | 1.04 | -.86 | .65 | -.87 |
| A Knew M/H | 1.13 | -1.18 | .94 | -1.01 |
| A Never H/M | .42 | -.44 | .01 | -.17 |
| A Never M/H | .38 | -.48 | .14 | -.19 |
| B Knew H/M | .84 | -.88 | .79 | -1.02 |
| B Knew M/H | .69 | -.80 | .53 | -.59 |
| B Never H/M | .22 | -.58 | .15 | -.53 |
| B Never M/H | .20 | -.79 | -.54 | .30 |
| SxLxO | | | | |
| A Complete H/M | .77 | -.71 | .47 | -.31 |
| A Complete M/H | .66 | -.83 | .39 | -.61 |
| A Shortened H/M | .69 | -.60 | .19 | -.74 |
| A Shortened M/H | .84 | -.83 | .69 | -.59 |
| B Complete H/M | .48 | -.48 | .33 | -.59 |
| B Complete M/H | .37 | -.66 | -.16 | -.15 |
| B Shortened H/M | .58 | -.98 | .61 | -.96 |
| B Shortened M/H | .51 | -.93 | .15 | -.14 |
| DxLxO | | | | |
| Knew Comp H/M | .99 | -.83 | .62 | -.82 |
| Knew Comp M/H | .84 | -.92 | .54 | -.95 |
| Knew Short H/M | .89 | -.91 | .82 | -1.08 |
| Knew Short M/H | .98 | -1.06 | .93 | -.65 |
| Never Comp H/M | .26 | -.36 | .18 | -.08 |
| Never Comp M/H | .20 | -.57 | -.31 | .19 |
| Never Short H/M | .38 | -.66 | -.02 | -.62 |

| | | | | |
|---|---|---|---|---|
| Never Short M/H | .37 | -.70 | -.08 | -.08 |
| SxDxU | | | | |
| A Knew Biased | 1.55 | -1.25 | 1.35 | -1.29 |
| A Knew N.B. | .62 | -.80 | .24 | -.59 |
| A Never Biased | .05 | -.24 | -.15 | .05 |
| A Never N.B. | .74 | -.68 | .31 | -.41 |
| B Knew Biased | 1.34 | -1.21 | 1.46 | -1.45 |
| B Knew N.B. | .19 | -.46 | -.14 | -.17 |
| B Never Biased | .04 | -.20 | -.05 | -.22 |
| B Never N.B. | .37 | -1.17 | -.33 | -.01 |
| SxLxU | | | | |
| A Comp Biased | .85 | -.77 | .59 | -.51 |
| A Complete N.B. | .58 | -.77 | .26 | -.40 |
| A Short Biased | .76 | -.72 | .60 | -.73 |
| A Short N.B. | .77 | -.70 | .29 | -.60 |
| B Comp Biased | .49 | -.52 | .53 | -.85 |
| B Complete N.B. | .36 | -.62 | -.36 | .11 |
| B Short Biased | .89 | -.89 | .87 | -.82 |
| B Short N.B. | .20 | -1.01 | -.11 | -.28 |
| SxOxU | | | | |
| A H/M Biased | .71 | -.65 | .37 | -.56 |
| A H/M Not Biased | .74 | -.65 | .29 | -.48 |
| A M/H Biased | .89 | -.84 | .82 | -.68 |
| A M/H N.B. | .61 | -.82 | .26 | -.52 |
| B H/M Biased | .76 | -.64 | .59 | -.80 |
| B H/M N.B. | .29 | -.82 | .36 | -.76 |
| B M/H Biased | .62 | -.78 | .82 | -.88 |

| | | | | |
|---|---|---|---|---|
| B M/H N.B. | .26 | -.81 | -.83 | .58 |
| LxOxU | | | | |
| Comp H/M Biased | .76 | -.61 | .58 | -.62 |
| Comp H/M N.B. | .49 | -.58 | .22 | -.28 |
| Comp M/H Biased | .58 | -.68 | .55 | -.74 |
| Comp M/H N.B. | .46 | -.82 | -.32 | -.01 |
| Short H/M Biased | .72 | -.68 | .38 | -.74 |
| Short H/M N.B. | .55 | -.90 | .42 | -.96 |
| Short M/H Biased | .93 | -.94 | 1.09 | -.82 |
| Short M/H N.B. | .42 | -.82 | -.24 | .08 |
| DxLxU | | | | |
| Knew Comp Bias | 1.34 | -1.16 | 1.22 | -1.40 |
| Knew Comp N.B. | .48 | -.59 | -.06 | -.37 |
| Knew Short Bias | 1.55 | -1.30 | 1.59 | -1.34 |
| Knew Short N.B. | .32 | -.67 | .16 | -.38 |
| Never Comp Bias | -.05 | -.13 | -.09 | .03 |
| Never Comp N.B. | .46 | -.81 | -.04 | -.08 |
| Never Short Bias | .10 | -.32 | -.12 | -.21 |
| Never Short N.B. | .65 | -1.04 | .02 | -.49 |
| Short M/H N.B. | .42 | -.82 | -.24 | .08 |
| DxOxU | | | | |
| Knew H/M Bias | 1.48 | -1.15 | 1.17 | -1.22 |
| Knew H/M N.B. | .40 | -.58 | .27 | -.67 |
| Knew M/H Bias | 1.41 | -1.31 | 1.64 | -1.52 |
| Knew M/H N.B. | .41 | -.68 | -.17 | -.08 |
| Never H/M Bias | -.01 | -.14 | -.21 | -.14 |
| Never H/M N.B. | .64 | -.89 | .37 | -.56 |

| | | | | |
|---|---|---|---|---|
| Never M/H Bias | .10 | -.31 | -.00 | -.04 |
| Never M/H N.B. | .47 | -.96 | -.39 | .15 |
| SxDxLxO | | | | |
| A Knew Comp H/M | 1.07 | -1.04 | .78 | -.67 |
| A Knew Comp M/H | 1.14 | -1.21 | .82 | -1.36 |
| A Knew Short H/M | 1.00 | -.68 | .52 | -1.08 |
| A Knew Short M/H | 1.11 | -1.16 | 1.06 | -.65 |
| A Never Comp H/M | .46 | -.37 | .16 | .05 |
| A Never Comp M/H | .18 | -.45 | -.05 | .14 |
| A Never Short H/M | .37 | -.51 | -.13 | -.39 |
| A Never Short M/H | .57 | -.50 | .33 | -.53 |
| B Knew Comp H/M | .90 | -.61 | .46 | -.97 |
| B Knew Comp M/H | .53 | -.64 | .26 | -.54 |
| B Knew Short H/M | .78 | -1.14 | 1.12 | -1.08 |
| B Knew Short M/H | .84 | -.96 | .79 | -.64 |
| B Never Comp H/M | .05 | -.36 | .21 | -.22 |
| B Never Comp M/H | .22 | -.68 | -.58 | .24 |
| B Never Short H/M | .38 | -.81 | .09 | -.84 |
| B Never Short M/H | .18 | -.90 | -.50 | -.36 |
| SxDxLxU | | | | |
| A Knew Comp Bias | 1.62 | -1.34 | 1.30 | -1.36 |
| A Knew Comp N.B. | .60 | -.92 | .31 | -.67 |
| A Knew Shrt Bias | 1.48 | -1.17 | 1.40 | -1.22 |
| A Knew Shrt N.B. | .63 | -.67 | .17 | -.51 |
| A Never Comp Bias | .08 | -.20 | -.10 | .33 |
| A Never Comp N.B. | .57 | -.63 | .21 | -.13 |
| A Never Shrt Bias | .03 | -.28 | -.20 | -.24 |

| | | | | |
|---|---|---|---|---|
| A Never Shrt N.B. | .92 | -.73 | .40 | -.68 |
| B Knew Comp Bias | 1.07 | -.99 | 1.14 | -1.44 |
| B Knew Comp N.B. | .36 | -.26 | -.42 | -.07 |
| B Knew Shrt Bias | 1.61 | -1.44 | 1.78 | -1.46 |
| B Knew Shrt N.B. | .01 | -.67 | .14 | -.26 |
| B Nvr Comp Bias | -.09 | -.06 | -.07 | -.26 |
| B Nvr Comp N.B. | .36 | -.98 | -.29 | .29 |
| B Nvr Shrt Bias | .17 | -.35 | -.03 | -.18 |
| B Nvr Shrt N.B. | .38 | -1.36 | -.36 | -.30 |
| SxDxOxU | | | | |
| A Knew H/M Bias | 1.45 | -1.04 | 1.07 | -1.09 |
| A Knew H/M N.B. | .63 | -.69 | .23 | -.66 |
| A Knew M/H Bias | 1.65 | -1.47 | 1.63 | -1.49 |
| A Knew M/H N.B. | .61 | -.91 | .26 | -.52 |
| A Nvr H/M Bias | -.03 | -.27 | -.32 | -.03 |
| A Nvr H/M N.B. | .86 | -.62 | .34 | -.31 |
| A Nvr M/H Bias | .13 | -.21 | .01 | .12 |
| A Nvr M/H N.B. | .62 | -.74 | .27 | -.51 |
| B Knew H/M Bias | 1.52 | -1.27 | 1.27 | -1.35 |
| B Knew H/M N.B. | .16 | -.48 | .31 | -.69 |
| B Knew M/H Bias | 1.16 | -1.16 | 1.65 | -1.55 |
| B Knew M/H N.B. | .21 | -.44 | -.59 | .36 |
| B Nvr H/M Bias | .01 | -.01 | -.09 | -.24 |
| B Nvr H/M N.B. | .42 | -1.16 | .40 | -.82 |
| B Nvr M/H Bias | .08 | -.40 | -.01 | -.20 |
| B Nvr M/H N.B. | .32 | -1.18 | -1.06 | .81 |
| SxLxOxU | | | | |

| | | | | |
|---|---|---|---|---|
| A Comp H/M Bias | .96 | -.80 | .63 | -.42 |
| A Comp H/M N.B. | .57 | -.62 | .30 | -.19 |
| A Comp M/H Bias | .73 | -.73 | .55 | -.60 |
| A Comp M/H N.B. | .60 | -.93 | .22 | -.61 |
| A Shrt H/M Bias | .46 | -.50 | .11 | -.70 |
| A Shrt H/M N.B. | .92 | -.69 | .27 | -.77 |
| A Shrt M/H Bias | 1.05 | -.95 | 1.08 | -.76 |
| A Shrt M/H N.B. | .63 | -.71 | .30 | -.42 |
| B Comp H/M Bias | .56 | -.43 | .52 | -.82 |
| B Comp H/M N.B. | .40 | -.54 | .15 | -.37 |
| B Comp M/H Bias | .43 | -.62 | .54 | -.88 |
| B Comp M/H N.B. | .32 | -.70 | -.86 | .58 |
| B Shrt H/M Bias | .97 | -.85 | .65 | -.78 |
| B Shrt H/M N.B. | .19 | -1.10 | .56 | -1.14 |
| B Shrt M/H Bias | .82 | -.94 | 1.09 | -.87 |
| B Shrt M/H N.B. | .20 | -.93 | -.79 | .58 |
| DxLxOxU | | | | |
| Knw Cmp H/M Bias | 1.48 | -1.14 | 1.09 | -1.20 |
| Knw Cmp H/M N.B. | .50 | -.52 | .15 | -.44 |
| Knw Cmp M/H Bias | 1.21 | -1.19 | 1.34 | -1.60 |
| Knw Cmp M/H N.B. | .47 | -.66 | -.26 | -.30 |
| Knw St H/M Bias | 1.49 | -1.17 | 1.25 | -1.25 |
| Knw St H/M N.B. | .29 | -.65 | .39 | -.91 |
| Knw St M/H Bias | 1.60 | -1.43 | 1.93 | -1.44 |
| Knw St M/H N.B. | .35 | -.69 | -.08 | .14 |
| Nvr Cmp H/M Bias | .04 | -.09 | .06 | -.04 |
| Nvr Cmp H/M N.B. | .47 | -.64 | .30 | -.12 |

| | | | | |
|---|---|---|---|---|
| Nvr Cmp M/H Bias | -.05 | -.16 | -.24 | .11 |
| Nvr Cmp M/H N.B. | .46 | -.97 | -.38 | .27 |
| Nvr St H/M Bias | -.06 | -.18 | -.48 | -.23 |
| Nvr St H/M N.B. | .82 | -1.14 | .44 | -1.01 |
| Nvr St M/H Bias | .26 | -.45 | .24 | -.19 |
| Nvr St M/H N.B. | .49 | -.95 | -.40 | .02 |
| SxDxLxOxI | | | | |
| A Knw Cmp H/M Bs | 1.57 | -1.25 | 1.17 | -.92 |
| A Knw Cmp H/M NB | .58 | -.84 | .39 | -.42 |
| A Knw Cmp M/H Bs | 1.67 | -1.42 | 1.42 | -1.80 |
| A Knw Cmp M/H NB | .62 | -1.00 | .23 | -.92 |
| A Knw Cmp H/M Bs | 1.34 | -.82 | .97 | -1.27 |
| A Knw St H/M NB | .67 | -.54 | .07 | -.89 |
| A Knw St M/H Bs | 1.63 | -1.51 | 1.84 | -1.18 |
| A Knw St M/H NB | .59 | -.82 | .28 | -.13 |
| A Nvr Cmp H/M Bs | .36 | -.35 | .10 | .07 |
| A Nvr Cmp H/M NB | .56 | -.40 | .21 | .04 |
| A Nvr Cmp M/H Bs | -.20 | -.04 | -.31 | .60 |
| A Nvr Cmp M/H NB | .57 | -.87 | .21 | -.31 |
| A Nvr St H/M Bs | -.41 | -.18 | -.74 | -.13 |
| A Nvr St H/M NB | 1.16 | -.85 | .48 | -.66 |
| A Nvr St M/H Bs | .47 | -.38 | .33 | -.35 |
| A Nvr St M/H NB | .67 | -.62 | .33 | -.71 |
| B Knw Cmp H/M Bs | 1.39 | -1.03 | 1.01 | -1.48 |
| B Knw Cmp H/M NB | .42 | -.20 | -.09 | -.46 |
| B Knw Cmp M/H Bs | .76 | -.96 | 1.27 | -1.40 |
| B Knw Cmp M/H NB | .31 | -.32 | -.75 | .31 |

| | | | | |
|---|---|---|---|---|
| B Knw Cmp H/M Bs | 1.65 | -1.52 | 1.53 | -1.23 |
| B Knw St H/M NB | -.09 | -.77 | .72 | -.93 |
| B Knw St M/H Bs | 1.58 | -1.35 | 2.03 | -1.70 |
| B Knw St M/H NB | .11 | -.57 | -.44 | .41 |
| B Nvr Cmp H/M Bs | -.27 | .17 | .03 | -.16 |
| B Nvr Cmp H/M NB | .38 | -.89 | .39 | -.28 |
| B Nvr Cmp M/H Bs | .10 | -.29 | -.18 | -.37 |
| B Nvr Cmp M/H NB | .34 | -1.08 | -.98 | .86 |
| B Nvr St H/M Bs | .30 | -.19 | -.22 | -.33 |
| B Nvr St H/M NB | .47 | -1.43 | .41 | -1.36 |
| B Nvr St M/H Bs | .06 | -.52 | .15 | -.04 |
| B Nvr St M/H NB | .30 | -1.29 | -1.14 | .76 |

## Appendix G

## OTHER RESULTS FROM THE INTENTIONAL-BIAS PHASE

A number of results from the Intentional-Bias phase were not of central importance to the hypotheses of the present study. For archival purposes, these effects are presented here. A significant multivariate main effect was found for Order ( $F$ (4,301)= 4.07, $p$ <.003). Examination of the associated univariates finds that only with FM was significance attained. For the 'hypothetical then memory' order, the mean score was -.65, for the 'memory then hypothetical' order, -.37. While this suggests a stronger hindsight bias for the former, the mean scores for both were significantly different from zero ( $p$ <.0001).

A significant multivariate main effect was found for Set ( $F$ (4,301)= 3.35, $p$ <.0106). Univariates were significant for TH and TM. For true labelled statements, a much stronger bias was found for 'list A'.

A significant multivariate interaction was found for Order by Set ( $F$ (4,301)= 5.88, $p$ <.0001). Univariates reveal that this effect was present with memory scores. Within the 'memory then hypothetical' order for TM, subjects who received 'list A' prior to the personality measures

produced a stronger bias ( p <.05) than those subjects who received 'list B'. For FM, subjects who received 'list B' produced a larger bias with the 'hypothetical then memory' order than with the 'memory then hypothetical' order ( p <.01).

Order by Length generated a significant multivariate interaction ( F (4,301)= 2.71, p <.03). The only significant univariate was for FM. Examining cell means, one finds that subjects who received 'shortened' instructions produced a stronger bias with the 'hypothetical then memory' order ( p <.01).

Understanding by Set was a significant multivariate interaction ( F (4,301)= 6.68, p <.0001). Univariates were significant for TM and FM. For both TM and FM, a significant bias was found with 'biased' instructions but the difference over 'not biased' instructions was significant only with list B ( p <.001). For TM, 'not biased' instructions produced a larger bias with list A ( p <.01).

Understanding by Order was a significant multivariate interaction ( F (4,301)= 10.27, p <.0001), reflected in univariates significant with memory scores. For both TM and FM, subjects who rated memory scored statements first produced a much stronger bias under 'biased' instructions ( p <.001). In addition, a much greater bias was found for the

'not biased' condition with the 'hypothetical then memory' order ( $p < .01$).

A significant 3-way multivariate interaction was found for Understanding by Length by Set ( $F$ (4,301)= 2.69, $p < .03$). This interaction was found only with TH. Examining cell means reveals that the 'shortened', 'not biased' and 'list B' cell appears to be the source of this interaction ( $p < .05$).

A second 3-way interaction was found for Understanding by Set by by Order ( $F$ (4,301)= 6.06, $p < .0001$). This 3-way interaction was found with memory scores. For TM, this interaction can be attributed to the 'not biased', 'list B' and 'memory then hypothetical' cell ( $p < .001$). The same cell but for 'list A' rather than for 'list B' is the source of the interaction for FM ( $p < .001$).

A significant multivariate Direction by Length interaction was found ( $F$ (4,301)=2.42, $p < .05$), but with a significant univariate for FM only. Subjects assigned to 'complete' instructions produced a larger bias with 'knew-it-all-along' over 'never-knew-that' instructions ( $p < .001$).

Finally, a significant 4-way multivariate interaction was detected ( $F$ (4,301)=3.26, $p < .01$). Independent variables involved were Set, Direction, Length, and Understanding. The

significant univariate was FM. For 'complete' instructions, a much larger hindsight bias was found for 'rerating' instructions with 'list A' than with 'list B' ( $p$ <.05), while biased instructions did not change. Ratings for 'knew'/'biased' differed significantly from 'knew'/'not biased' for 'list A', and from 'not biased'/'never' ( $p$ <.01), 'biased'/'never' ( $p$ <.01), and 'not biased'/'knew' ( $p$ <.001). For 'shortened' instructions, a similar pattern was found. For 'list A', 'knew'/'not biased' differed from 'knew'/'biased' and for 'list B', from 'knew'/'not biased','never'/'not biased' and 'never'/'biased' ( $p$ <.01).

# Appendix H

## STATISTICS FOR THE PERSONALITY MEASURES

Table H-1

Backward Elimination Procedure for
True-Hypothetical Scores,
No Instruction Subjects

| Variable | Partial Sum of Squares | Beta Weight | F | p |
|---|---|---|---|---|
| All variables included in the model, R square = .0836 | | | | |
| M.C. | 2.8145 | .0192 | 5.68 | .02 |
| T.S.B.I. | .1730 | .0010 | .03 | .8522 |
| S.M. | .8301 | -.0161 | 1.68 | .1998 |
| Variable T.S.B.I. Removed, R square = .0831 | | | | |
| M.C. | 2.8511 | -.0192 | 5.84 | .0183 |
| S.M. | .8131 | -.0157 | 1.67 | .2012 |
| Variable S.M. Removed, R square = .0607 | | | | |
| M.C. | 2.1970 | .0161 | 4.46 | .0384 |

Note: 12 observations deleted due to missing values
M.C. = Marlowe-Crowne Scale
T.S.B.I. = Texas Social Behavior Inventory
S.M. = Self-monitoring Scale

Table H-2

Backward Elimination Procedure for
False-Hypothetical Scores,
No Instruction Subjects

| Variable | Partial Sum of Squares | Beta Weight | F | p |
|---|---|---|---|---|
| All variables included in the model, R square = .0895 | | | | |
| M.C. | .0319 | -.0020 | .04 | .8455 |
| T.S.B.I. | 5.3283 | -.0192 | 6.39 | .0138 |
| S.M. | .4962 | .0124 | .60 | .4431 |
| Variable M.C. Removed, R square = .0890 | | | | |
| T.S.B.I. | 5.3969 | -.0193 | 6.57 | .0126 |
| S.M. | .4653 | .0115 | .57 | .4544 |
| Variable S.M. Removed, R square = .0814 | | | | |
| T.S.B.I. | 4.9937 | -.0182 | 6.12 | .0159 |

Note: 12 observations deleted due to missing values.
M.C. = Marlowe-Crowne scale
T.S.B.I. = Texas Social Behavior Inventory
S.M. = Self-monitoring scale

Table H-3

Backward Elimination Procedure for
True-Memory Scores,
No Instruction Subjects

| Variable | Partial Sum of Squares | Beta Weight | F | p |
|---|---|---|---|---|
| All variables included in the model, R square = .0408 | | | | |
| M.C. | .0468 | .0025 | .15 | .6982 |
| T.S.B.I. | .0112 | .0009 | .04 | .8491 |
| S.M. | .8609 | -.0166 | 2.79 | .0996 |
| Variable T.S.B.I. Removed, R square = .0402 | | | | |
| M.C. | .0497 | .0025 | .16 | .6873 |
| S.M. | .8510 | -.0163 | 2.80 | .0990 |
| Variable S.M. Removed, R square = .0379 | | | | |
| M.C. | .8047 | -.0151 | 2.68 | .1063 |

Note: 13 observations deleted due to missing values.
M.C. = Marlowe-Crowne scale
T.S.B.I. = Texas Social Behavior Inventory
S.M. = Self-monitoring scale

Table H-4

Backward Elimination Procedure for
False-Hypothetical Scores,
No Instruction Subjects

| Variable | Partial Sum of Squares | Beta Weight | F | p |
|---|---|---|---|---|
| All variables included in the model, R square = .0455 | | | | |
| M.C. | .2160 | .0053 | 1.12 | .2932 |
| T.S.B.I. | .0688 | -.0022 | .36 | .5518 |
| S.M. | .4072 | -.0114 | 2.12 | .1505 |
| Variable T.S.B.I. Removed, R square = .0403 | | | | |
| M.C. | .2027 | .0051 | 1.06 | .3061 |
| S.M. | .4713 | -.0121 | 2.47 | .1205 |
| Variable M.C. Removed, R square = .0251 | | | | |
| S.M. | .3335 | -.0097 | 1.75 | .1905 |

Note: 13 observations deleted due to missing values.
M.C. = Marlowe-Crowne scale
T.S.B.I. = Texas Social Behavior Inventory
S.M. = Self-monitoring scale

Table H-5

Summary Statistics for the Personality Measures
No Instruction Subjects

| Variable | M | Var | Range |
|---|---|---|---|
| M.C. | 103.48 | 135.98 | 51-125 |
| T.S.B.I. | 90.85 | 206.13 | 60-127 |
| S.M. | 54.71 | 52.01 | 38-71 |

Note: M.C. = Marlowe-Crowne scale
T.S.B.I. = Texas Social Behavior Inventory
S.M. = Self-monitoring scale

Table H-6

Backward Elimination Procedure for
True-Hypothetical Scores,
Debiasing Subsample

| Variable | Partial Sum of Squares | Beta Weight | F | p |
|---|---|---|---|---|
| All variables included in the model, R square = .0426 | | | | |
| M.C. | .0622 | .0032 | .14 | .7118 |
| T.S.B.I. | .3460 | .0053 | .77 | .3847 |
| S.M. | .9866 | -.0223 | 2.19 | .1445 |
| Variable M.C. Removed, R square = .0404 | | | | |
| T.S.B.I. | .3309 | .0052 | .74 | .3918 |
| S.M. | .9253 | -.0210 | 2.08 | .1544 |
| Variable T.S.B.I. Removed, R square = .0285 | | | | |
| S.M. | .7927 | -.0193 | 1.79 | .1859 |

Note: 18 observations deleted due to missing values
M.C. = Marlowe-Crowne Scale
T.S.B.I. = Texas Social Behavior Inventory
S.M. = Self-monitoring Scale

Table H-7

Backward Elimination Procedure for
False-Hypothetical Scores,
Debiasing Subsample

| Variable | Partial Sum of Squares | Beta Weight | F | p |
|---|---|---|---|---|
| All variables included in the model, R square = .0617 | | | | |
| M.C. | .1547 | .0051 | .34 | .5622 |
| T.S.B.I. | 1.5284 | .0111 | 3.36 | .0719 |
| S.M. | .0099 | .0022 | .02 | .8831 |
| Variable S.M. Removed, R square = .0614 | | | | |
| T.S.B.I. | 1.6004 | .0113 | 3.57 | .0635 |
| M.C. | .1815 | .0054 | .41 | .5267 |
| Variable M.C. Removed, R square = .0551 | | | | |
| T.S.B.I. | 1.5758 | .0112 | 3.55 | .0642 |

Note: 18 observations deleted due to missing values.
M.C.= Marlowe-Crowne scale
T.S.B.I. = Texas Social Behavior Inventory
S.M. = Self-monitoring scale

Table H-8

Backward Elimination Procedure for
True-Memory Scores,
Debiasing Subsample

| Variable | Partial Sum of Squares | Beta Weight | F | p |
|---|---|---|---|---|
| All variables included in the model, R square = .0258 | | | | |
| M.C. | .0102 | -.0013 | .05 | .8264 |
| T.S.B.I. | .3104 | -.0050 | 1.48 | .2287 |
| S.M. | .0435 | .0047 | .21 | .6506 |
| Variable M.C. Removed, R square = .0250 | | | | |
| T.S.B.I. | .3050 | -.0050 | 1.48 | .2289 |
| S.M. | .0364 | .0042 | .18 | .6759 |
| Variable S.M. Removed, R square = .0221 | | | | |
| T.S.B.I. | .2818 | -.0047 | 1.38 | .2440 |

Note: 18 observations deleted due to missing values.
M.C. = Marlowe-Crowne scale
T.S.B.I. = Texas Social Behavior Inventory
S.M. = Self-monitoring scale

Table H-9

Backward Elimination Procedure for
False-Memory Scores,
Debiasing Subsample

| Variable | Partial Sum of Squares | Beta Weight | F | p |
|---|---|---|---|---|
| All variables included in the model, R square = .0531 | | | | |
| M.C. | .0568 | -.0031 | .21 | .6456 |
| T.S.B.I. | .3133 | .0050 | 1.18 | .2822 |
| S.M. | .4263 | .0146 | 1.60 | .2105 |
| Variable M.C. Removed, R square = .0496 | | | | |
| T.S.B.I. | .3294 | .0052 | 1.25 | .2671 |
| S.M. | .3786 | .0135 | 1.44 | .2345 |
| Variable T.S.B.I. Removed, R square = .0298 | | | | |
| S.M. | .4933 | .0152 | 1.87 | .1764 |

Note: 18 observations deleted due to missing values.
M.C. = Marlowe-Crowne scale
T.S.B.I. = Texas Social Behavior Inventory
S.M. = Self-monitoring scale

Table H-10

Summary Statistics for the Personality Measures, Debiasing Subsample

| Variable | M | Var | Range |
|---|---|---|---|
| M.C. | 102.13 | 110.04 | 70-135 |
| T.S.B.I. | 98.73 | 189.06 | 71-121 |
| S.M. | 53.97 | 33.41 | 43-68 |

Note: M.C. = Marlowe-Crowne scale
T.S.B.I. = Texas Social Behavior Inventory
S.M. = Self-monitoring scale