

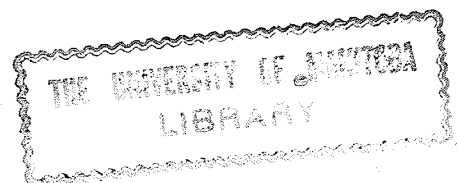
INDUCEMENT OF EXPECTANCY AND SET OF SUBJECTS
AS DETERMINANTS OF SUBJECTS' RESPONSES
IN EXPERIMENTER EXPECTANCY RESEARCH

A Dissertation
Presented to
The Faculty of Graduate Studies and Research
University of Manitoba

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
Ronald Walter Johnson

October 1970



ACKNOWLEDGEMENT

I wish to express my sincere appreciation to Dr. Marion Aftanas and to Dr. Robert Altemeyer for their constructive criticisms both during the planning of the project and during the preparation of the manuscript. To Dr. Robert Rosenthal I owe my interest in the subject matter of this dissertation. He has served as an inspiration in my small attempts at adding to the wealth of scientific information that he has collected.

Especially I should like to thank Dr. John Adair who not only aided in countless ways with this dissertation, but who has so frequently acted as a wise and selfless counsel to me throughout the course of my studies.

Finally I should like to thank my wife Betty who, in addition to providing the usual psychological support, aided in the planning of this investigation, who acted as the "additional investigator" in collecting the data, who helped with the statistical analyses, and who did most of the typing.

ABSTRACT

INDUCEMENT OF EXPECTANCY AND SET OF SUBJECTS AS DETERMINANTS OF SUBJECTS' RESPONSES IN EXPERIMENTER EXPECTANCY RESEARCH

by

Ronald Walter Johnson

The purpose of the present study was to investigate a model of experimenter expectancy which postulates that expectancy effects increase with an increasing set to transmit hypothesis-related information by the experimenter, and with an increasing set to receive that information by the subject. In an attempt to clarify two methodological criticisms of the expectancy literature a control was exercised for observer-recorder error and intentionality of the experimenter was manipulated.

Eighteen experimenters were randomly assigned to three conditions of expectancy inducement. In one condition the hypothesis was simply stated to the experimenters, in a second condition the principal investigator role-played great outcome concern, and in a third condition the experimenters were asked to actively manipulate subjects' responses. Two hundred sixteen subjects were randomly assigned to three

conditions of evaluation apprehension which attempted to arouse evaluation apprehension to three different levels. The method of randomization allowed for a check on effects of early-testing versus late-testing of subjects.

Where the principal investigator showed outcome concern a significant interaction between expectancy and evaluation apprehension, across the six trials of a marble-dropping task, was obtained. Where the experimenters were simply told the hypothesis no expectancy effects were noted. However, where the experimenters attempted to actively manipulate results significant effects opposite to their expectancies were observed. An analysis of tape-recordings of the experimental sessions revealed that intentional-inducement experimenters made far more verbalizations than did other experimenters.

The results of the investigation were interpreted as indicating that a minimum of both experimenter outcome concern and subject performance concern must be present for expectancy mediation. It was postulated that the reversed expectancy effects for subjects tested by intentional-inducement experimenters might have been due to subjects reacting against strong cues transmitted by the experimenters.

TABLE OF CONTENTS

	Page
List of Tables	v
List of Figures	vi
Chapter	
I. Introduction	1
II. Method	22
III. Results	32
IV. Discussion	53
References	66
Appendices	72

LIST OF TABLES

Table	Page
1. Analysis of Variance on Numbers of Marbles Dropped per Minute by Subjects of Low-Inducement Experimenters	35
2. Analysis of Variance on Numbers of Marbles Dropped per Minute by Subjects of High-Inducement Experimenters	38
3. Analysis of Variance on Numbers of Marbles Dropped per Minute by Subjects of Intentional-Inducement Experimenters	43

LIST OF FIGURES

Figure	Page
1. Mean responses for subjects tested by low-inducement experimenters	36
2A. Mean responses for high evaluation apprehension subjects tested by high-inducement experimenters	40
2B. Mean responses for medium evaluation apprehension subjects tested by high-inducement experimenters	41
2C. Mean responses for low evaluation apprehension subjects tested by high-inducement experimenters	42
3. Mean responses for subjects tested by intentional-inducement experimenters	45

CHAPTER I

INTRODUCTION

Experimenter expectancy effects refer to the degree to which "an experimenter's expectations can affect the data actually obtained in his research [Rosenthal, 1964, p. 97]." It has been assumed by Rosenthal that "the experimenter has certain expectations about the relationship or lack of relationship between the selected variables and certain other variables [1964, p. 97]." Not only does the experimenter frequently have a subjective prediction as to what form his to-be-collected data will take, he may also have certain investments in the nature of that data. He may, for example, desire the discovery of significant relationships between investigated variables for no other reason than his concern with the advancement of scientific knowledge.

The early experimental demonstrations of Rosenthal's hypothesis of the prevalence of these experimenter expectancy effects have been well documented (Rosenthal, 1963; Rosenthal and Fode, 1963a) and few psychologists do not have at least a passing acquaintance with Rosenthal's work. The first experiment was paradigmatic and will be described to illustrate the most common technique which has been used in expectancy research.

In order to conduct experiments on experimenter expectancy, Rosenthal and Fode (1963a) presented 57 photographs of faces cut from a weekly newsmagazine to 104 subjects instructed to rate each photograph with reference to the degree to which the individual photographed was experiencing success or failure. The rating scale used by the subjects ran from -10 for very unsuccessful to +10 for very successful. From the 57 photographs Rosenthal and Fode selected 20 which had been judged by their subjects to be, on the average, "neutral" (either -1 or +1) as to their stimulus value. Next they created two groups of experimenters who were required to have subjects rate the 20 photographs mentioned above on the rating scale. Each experimenter was told that his task was to attempt a replication of certain "well-established experimental findings". One group of experimenters was told to expect an average of -5 rating on the neutral photographs, the other group was told to expect an average +5 rating. The difference between the data obtained by the two groups of experimenters was statistically reliable; those experimenters who expected -5 ratings obtained significantly more negative ratings than did those experimenters expecting +5 ratings.

A new variable in psychological research frequently stimulates considerable excitement as investigators attempt to observe relationships in a variety of situations. Experimenter expectancy has been no different, as principal

investigators sought and found the phenomenon to be present in such divergent places as the animal laboratory (Rosenthal and Fode, 1963b; Rosenthal and Lawson, 1964), the clinic (Marwit and Marcia, 1967), the sensory deprivation chamber (Raffetto, 1968), and even the classroom (Rosenthal and Jacobson, 1968).¹

Criticism of experimenter expectancy

Even though there has been a proliferation of attempts at demonstrating the pervasiveness of experimenter expectancy, a decade of intensive research has not led to any real understanding of the complex interactions between the experimenter and the subject which cause some experimenters to bias the results of their research. Controversy has arisen over the pervasiveness of experimenter expectancy effects in various experimental settings. As well, the very existence of the effects in any setting of importance to behavioural scientists has been questioned by critics of the literature. A number of recent "negative" findings have raised the question as to whether the effect is really worthy of investigation as a pervasive phenomenon of appreciable impact, or whether it is merely an artifact of a few poorly designed studies.

As the initial enthusiasm dampened various investigators have failed to replicate earlier findings (Bar-

¹Several reviews of this early literature on experimenter expectancy have been published (Barber and Silver, 1968a; Friedman, 1967; Rosenthal, 1964; Rosenthal, 1966; Rosenthal, 1967; Rosenthal, 1968; Rosenthal, 1969).

ber, Caverley, Forgiione, McPeake, Chaves, and Brown, 1969; Silver, 1968; Wessler and Strauss, 1968). Barber and Silver (1968a), in reviewing 31 studies on the expectancy effect, have argued that 19 of these studies failed to demonstrate the effect, or do not "lend themselves to clear-cut conclusions [p. 37]." Rosenthal (1968) has questioned the legitimacy of Barber and Silver's interpretation that these studies do not demonstrate experimenter bias. The controversy (Barber and Silver, 1968b) centers around interpretation of "significance levels", "probability pyramiding", and the significance of "runs of experiments". However, the crucial issue concerns not these studies, which to some degree show "negative" results, but those studies, including the twelve others reviewed by Barber and Silver, which indicate statistically significant results irrespective of who interprets the statistical manipulations.

Bias not affecting subjects' responses versus experimenter expectancy. Barber and Silver (1968a; 1968b) point out that bias effects may be mediated by two sets of modes. One set (hence to be referred to as Type A) includes those modes that do not affect the subjects' responses; i. e., "unintentional or intentional misjudgment or misrecording of the responses on the part of the experimenter and fabrication of the data by the experimenter [1968a, p. 187]." The other set (hence to be referred to as Type B) includes those modes which actually affect the responses of the subjects, whether or not these effects

were intended by the experimenter.

Rosenthal (1969) has subdivided the first set of modes into observer error, recorder error, interpreter error, and intentional error. Effects mediated by those modes not affecting subjects' responses are artifacts of behavioural research which are not experimenter expectancy effects in the definition used by Rosenthal and by Barber and Silver. This definition restricts experimenter bias to effects influencing subjects' responses through "unintentional paralinguistic and kinesic cues". In other words, only where bias has been demonstrated to have been mediated by an unintended Type B mode can one argue that experimenter expectancy has been demonstrated. In their reply to Rosenthal, Barber and Silver (1968b) suggest that only two studies have demonstrated bias where the possibility of "misjudging vague criteria or misrecording responses" was ruled out (p. 59).

In one of these studies, Masling (1965) had graduate students administer the Rorschach test under conditions where the experimenters either expected many human responses or expected many animal responses. All testing was tape-recorded and scoring was done by independent observers from the tapes, so that the significant results cannot be attributed to observer or recorder error.

In the other study, Adair and Epstein (1968) obtained a bias effect in a condition where the subjects received only taped instructions on the person-perception

task. The taped instructions were obtained by means of a concealed recorder which taped the instructions given in an earlier part of the study where a significant bias effect was also obtained. It should be noted that in both conditions the subjects recorded their own responses, eliminating the possibility of the experimenter's committing a recording error.

Barber and Silver thus suggest that experimenter expectancy has been demonstrated in only two studies which controlled for observer or recorder bias. However, three further studies have since either exercised this control or have attempted to separate the two types of effects. Johnson (1970) controlled for a Type A effect, and demonstrated a significant expectancy main effect. Johnson used a marble-dropping task where the main dependent variable was a change in response rate as measured by difference scores between the first and last trials, computed for each subject. The data were recorded, not by the experimenters, but by the principal investigator who worked in an adjacent room and who was not aware of the conditions of the subjects as they were being tested. Dusek (1970) used the same task as Johnson, but his subjects were school-age children rather than college sophomores. Dusek found a significant bias effect for girls but not for boys.

Johnson and Adair (1970) used a word association task with latency of response as the dependent measure. The experimenters in this study observed and recorded the

data. However, continuous tape recordings were made of the experimental sessions allowing a later re-recording of the latencies by an independent observer who was kept blind as to the experimental conditions. While some observer error was found to be attributable to the experimenters, twice as much variance was accounted for by the main effect of experimenter expectancy.

For the majority of the studies, however, the Barber and Silver criticism is still valid and it is possible that in some of these experiments which have been cited to support the expectancy hypothesis, the effect may have been due to an artifact where subjects' responses were not affected. However, on the basis of the five studies reviewed above, which obtained significant effects even when this artifact was controlled or measured, it is hypothesized that the experimenter expectancy effect will be a significant source of variation where the subjects' responses, independent of observational or recording error, are affected. This prediction is necessitated by the observation that the only published experiments which have attempted to test the expectancy hypothesis in a situation controlling for observer and recorder error have demonstrated the phenomenon. On the other hand, the possibility that negative findings have been obtained but not reported cannot be ruled out.

Variables necessary for the experimenter expectancy effect

Experimenter expectancy research, as has been

pointed out by Levy (1969), suffers from a paucity of theory with which one can account for the increasing volume of conflicting empirical observations. Not only has the issue concerning which modes mediate bias not been resolved, but also the related variables of the phenomenon have not been established. It is important to establish under what conditions one might predict the presence of the phenomenon.

One way of viewing the expectancy experiment is as a two-person interaction situation where one person (the experimenter) communicates certain information to the other person (the subject). The experimenter has knowledge of the manner in which it is predicted the subject will perform. Only when the experimenter communicates this information, or a part of this information, to the subject, and only when the subject acts on this information, can the results be biased in the direction of the hypothesis; provided that there is no observer or other artifactual error.

Numerous attempts have been made to relate certain variables with the presence or absence of the expectancy phenomenon. If experimenter expectancy is mediated by informational communication from the experimenter to the subject, two broad variables need be considered for the presence or absence of the expectancy effect. These two variables, one concerning the experimenter's set to communicate information, the other concerning the subject's set to receive that information, can then account for a number of related variables and "explain" much of the re-

search on experimenter expectancy.

In order for an experimenter to communicate an expectancy to a subject, both the communicator and the receiver must be set for that communication process. It may well be that neither a set, or motivation, to send information, nor a set, or motivation, to receive that information is alone a sufficient force to mediate an expectancy. It is the intention of the present investigation to explore these two variables simultaneously in order to test the notion that the expectancy effect increases as set to transmit and set to receive are incremented to a maximum.

Methods of inducement. A crucial variable of the expectancy phenomenon, which could influence experimenters to cue subjects as to how to respond, is the method of inducing the expectancy. As has been pointed out by Barber and Silver (1968a), this variable has not been systematically investigated. The first step in any bias experiment is to induce an expectancy in the experimenter. How this is done may determine whether or not the effect, if present at all, will control a significant amount of the total variance. It is possible, for example, that a principal investigator, himself expecting to produce the effect, will behave differently in inducing his experimenters than will an investigator less convinced that the effect is an important one, and in fact possibly committed to obtaining negative results. When a principal investigator

instructs his experimenters he is in fact role-playing. He is deceiving these experimenters, and his effectiveness will depend upon whether his cover story is believed, upon whether the experimenters perceive the task as being important, upon whether they wish "to produce" for this particular investigator, etc. It is possible that with a thorough investigation of the means of inducing bias many of the questions on the pervasiveness of the phenomenon may be answered.

There have, nonetheless, been several reported instances where experimenters obtained different results as a function of differences in inducement. Rosenthal, Persinger, Mulry, Vikan-Kline and Grothe (1964) found that male experimenters who were made more conscious of the importance of experimental procedures obtained ratings of persons on the person-perception task as significantly less successful than did experimenters not made procedure-conscious. The authors suggest that perhaps these experimenters were made to feel their importance as data-collectors and this aided mediation of the effect.

Adler (1968) found that when scientific methodology was stressed for some experimenters and outcome was stressed for others the interaction between experimenter set and expectancy closely approached significance ($p < .06$), indicating that outcome-oriented experimenters obtained greater differences in responses from high and low expectancy subjects than did methodology-oriented experimenters.

Rosenthal, Persinger, Vikan-Kline, and Mulry (1963) trained experimenters for a bias study and in turn had these now experienced experimenters instruct new assistants. They found that the greater a given experimenter's biasing effect on his own subjects, the greater the effect of that experimenter's assistants on their subjects. It would seem that possibly a principal investigator's enthusiasm for his research may transfer to the experimenters.

Johnson and Adair (1970) attempted to manipulate two methods of inducement of expectancy where for one-half of the experimenters the principal investigator role-played greater concern for the importance of the research and for the outcome than he did for the other half of the experimenters. While the expectancy main effect reached conventional statistical significance ($p < .05$), the predicted interaction between expectancy and inducement did not achieve this criterion in the overall analysis of variance, having an associated probability level of only .10. The investigators admit that the inducement manipulation was not completely effective, as the principal investigator attempted to spend equal time with both groups of experimenters. As he thus spent much time with all experimenters, it is probable that the low level of inducement was closer to the high level of inducement than was desirable. The main difference in role-playing concerned whether or not the principal investigator checked on the data as it was being collected; i. e., on whether or not he pointedly asked to see the early data

returns. It is thus hypothesized that a clearer differentiation between the methods of inducement, where less enthusiasm for the research and its importance is communicated in a low-inducement condition, will produce a significant difference in the expectancy for high-inducement and low-inducement conditions.

Barber and Silver (1968a; 1968b) raise an additional criticism of previous research by questioning the intentionality of the expectancy effect. They question whether or not the experimenter, who somehow affects subjects' responses in the direction of his hypothesis, could have done so unintentionally, as is the criterion set forth by Rosenthal. Intentionality has long been a difficult issue in psychological research and a completely satisfactory answer to the question of whether or not a particular observed behaviour is intentional has never been found. Nonetheless, the question has been raised with reference to experimenter bias. To date no experiments on experimenter bias have been conducted which attempted to manipulate intentionality. However, it is proposed that there is a monotonic relationship between the degree to which the experimenter is committed to the substantiation of the hypothesized relationship and the degree of the expectancy effect. The highest level of inducement that could be achieved would be one where the experimenter is induced to attempt actively to manipulate his results. Thus the level of inducement which should produce the greatest ex-

pectancy effect would be one where the experimenter is completely briefed, prior to testing his subjects, and told to try to get certain results without, of course, telling the subject how to perform, without using verbal reinforcement, or without "fudging" the data. In other words, the highest level of inducement would be one where the experimenter attempts to communicate, nonverbally, to the subject the desired responses. It is thus hypothesized that the expectancy effect will be greater under these conditions than under the low-inducement or the high-inducement conditions.

Early data returns. Rosenthal (1966) has pointed out that the effects of hypothesis confirmation or disconfirmation by early-tested subjects affect the experimenter's behaviour with subjects tested later in the experiment. Rosenthal has presented evidence to suggest that when early-tested subjects confirm the hypothesis, experimenters tend to bias later subjects; but, when early-tested subjects disconfirm the hypothesis, the experimenter may change his expectancies and thus bias later-tested subjects to different degrees, or in different directions. This interpretation assumes that the experimenter is aware of how the subjects are performing as each subject is tested. In not all experimental situations, however, is feedback as to the subject's performance immediate. It seems important, therefore, to check on the performance of early-tested subjects versus later-tested subjects in the situ-

ation where minimal feedback, as to each subject's performance, is given the experimenter. One may assume that in this situation few differences in performances of subjects tested at various times are to be expected. However, empirical evidence has not been collected for this situation. In addition, there is a fairly wide variability in the number of subjects tested by each experimenter from one expectancy study to another. As a consequence it seems necessary to measure the effects of experimenter expectancy at different stages in the experiment. If it is found that there are differences in expectancy effects due to time of testing, the number of subjects tested by a single experimenter could be important in the mediation of the expectancy phenomenon.

Task ambiguity. One partially confirmed hypothesis suggests that as ambiguity increases in the experimental situation the potentiality of the bias effect also increases. Weiss (1969) attempted to manipulate ambiguity by changing exposure time for a number of dots exposed tachistoscopically. Although by manipulating exposure time the author also manipulated the structure of the stimulus, there is partial confirmation, for female subjects, for the hypothesis that experimenter bias increases with increasing stimulus ambiguity.

Shames and Adair (1967) used two tasks, the person-perception task and a numerosity-estimation task, to examine the generality of the experimenter bias effect. Their

conclusion, based on significant effects for the former but not the latter task, suggests that with the ambiguous person-perception task the subject must look to the experimenter for cues as to how to respond. In contrast, for the numerosity-estimation task the subject may perceive the situation as being relatively straightforward and experience no conflict as to how to proceed.

In two studies using a variety of tasks, Wessler found no significant expectancy effects for reaction time (Wessler, 1968) and no significant effects for the person-perception task, line estimation, or dot tapping (Wessler, 1969). In the second study, however, using a more liberal criterion for expectancy effects, Wessler found that subjects who perceived the photographs of the person-perception task to be in line with the experimenters' expectancies also produced line estimations co-ordinate with the experimenters' expectancies, but they showed no differences in dot tapping. Wessler concludes that the more obvious the correct response is to the subject, the less the probability that he will be influenced by experimenter bias.

If the expectancy effect increases with task ambiguity, as seems to be the case, it is probably due to the fact that the subject is in greater need of knowledge as to how to respond in an ambiguous situation, and therefore looks to the experimenter for appropriate cues.

Set of subjects. If it is correct to view the expectancy experiment as an interaction situation where

the experimenter communicates information to the subject, the expectancy effect should be facilitated when the subject is set to receive cues from the experimenter as to how to respond. An entirely different source of cue-seeking behaviour from that effected by task ambiguity may be effected by manipulating the set of the subject prior to his entering the experimental room.

Rosenberg (1969) has suggested that the subject "who is possessed of a concern over evaluation may well be more closely attuned to such indirect communication" as information emitted by an experimenter (p. 322). The subject, according to this notion, is more or less aroused to the impression that he is somehow being evaluated by the "Psychologist". "Evaluation apprehension" is defined by Rosenberg (1965, p. 29) as:

an active, anxiety-toned concern that he win a positive evaluation from the experimenter, or at least that he provide no grounds for a negative one. Personality variables will have some bearing upon the extent to which this pattern of apprehension develops. But equally important are various aspects of the experimental design such as the experimenter's explanatory "pitch", the types of measures used, and the experimental manipulations themselves.

Where his evaluation apprehension has been greatly aroused, the subject may look to the experimenter for cues as to how to perform, and for feedback as to how he is performing. In his attempt to earn a favourable evaluation he may thus be affected by the behaviour of the experimenter. If, in fact, the experimenter has certain expectations concerning the subject's performance, these expectations may very

well be communicated to the subject and consequently fulfilled.

Rosenberg (1969) has found support for this contention that "evaluation apprehension" is an important mediator of experimenter expectancy. In a series of experiments apprehension was aroused by written communications read by the subjects prior to being tested on the person-perception task. In the first of this series, a predicted and significant interaction ($p < .002$) was obtained between high and low evaluation apprehension conditions and +5 and -5 expectancy groups. A second experiment failed to replicate the results of the first, and Rosenberg suggests that this was due to a failure of the experimental procedures leading to the arousal of undue suspicion about the experiment. In the third experiment taped instructions, which were designed to be "shaded" in a positive ("success" stressing) or negative ("failure" stressing) direction, were obtained. When these taped instructions were played back to subjects, in low evaluation apprehension, control evaluation apprehension, and high evaluation apprehension conditions, a significant linear trend for the predicted order of increasing bias effects with increasing arousal of evaluation apprehension was obtained.

Rosenberg, in addition to having problems with suspicion in the second experiment of this series, observed manipulation difficulties in the third experiment. He reports that 13 (out of the original 216 female subjects) subjects were eliminated from the analysis because of aware-

ness of the experiment's purpose, and 22 were eliminated because they did not correctly rate the "evaluation apprehension" communications as either "anxiety arousing" or "reassuring" on a post-experimental questionnaire (p. 334). Although Rosenberg does not elaborate on the reasons for these difficulties, it would seem desirable to increase the credulity of the manipulations as experienced by the subject. In Rosenberg's experiments, subjects read communications which either led them to believe that poor performance on the person-perception task was associated with psychopathology, or that they were part of a control or standardization group. In the former case it may have been that some subjects simply did not believe this stated relationship, or else they may have correctly decided that this communication was simply being presented to see if they would believe it, and consequently be affected by it in their task performance.

A better test of the relationship between evaluation apprehension and experimenter expectancy could be made by having a group of subjects, in addition to reading "arousal" communications, engage in a task which would give plausibility to the contention that they are being evaluated on an important dimension. For example, in a situation where a subject is told that intelligence is related to performance on a specified task, he should be more inclined to infer that his intelligence is being evaluated on that task if he completes, during the same experimental session,

a standard ability test containing considerable face validity. It is thus hypothesized that the expectancy effect will be greater when a subject has read a communication designed to arouse his evaluation apprehension than when he has read a non-arousing communication; and, that the experimental effects will be greater still when he has performed a task which adds credulity to the apprehension communication.

Statement of the problem

Experimenter expectancy research has not led to a real understanding of the expectancy phenomenon. Certain methodological problems, in particular the separation of observer error from effects on subjects' responses, have tended to obscure the real issues. Experimenter expectancy is here viewed as an interaction situation where the experimenter transmits information about expected performance to the subject. Crucial to the mediation of the expectancy is the method of inducing the expectancy or bias in the experimenter, and the set of the subjects prior to the experimental testing.

In order to test the contention that the expectancy effect increases as set to transmit and set to receive information about performance are incremented, the following experimental investigation was undertaken. This investigation manipulated three methods of inducing expectancies; one method in which the principal investigator simply stated certain expected responses to the experimenters, a second method in which the principal investigator role-played

great concern about the importance of the outcome of the work, and in a third situation the principal investigator requested that the experimenters purposely attempt to manipulate their results. Secondly, this investigation manipulated three levels of one type of set of subjects, that variable being evaluation apprehension. As well, a check was made on the effects of early-tested versus late-tested subjects in a situation where few feedback cues as to the subjects' performances were given the experimenters. The main dependent variable concerned changes in response rates over time. A secondary dependent variable concerned absolute response rates of subjects tested under different conditions.

Several predictions, related to the two induced expectancies (high increase in performance and low increase in performance), were made concerning the performances of the subjects. These predictions were:

1. There would be a significant expectancy effect, i. e., subjects in the high expectancy conditions would show greater improvement in the rate of performance across trials than would subjects in the low expectancy conditions.

- 2A. The expectancy effect would be dependent upon the method of inducement such that the differences between improvement scores for high and low expectancy subjects would be more positive in the high inducement conditions than in the low inducement conditions.

2B. The expectancy effect would be dependent upon the method of inducement such that the differences between improvement scores for high and low expectancy subjects would be more positive in the intentional inducement conditions than in either the high or low inducement conditions.

3A. There would be a significant interaction between set of subjects and expectancy such that the difference between improvement scores for high and low expectancy subjects would be more positive in the medium evaluation apprehension conditions than in the low evaluation apprehension conditions.

3B. There would be a significant interaction between set of subjects and expectancy such that the difference between improvement scores for high and low expectancy subjects would be more positive in the high evaluation apprehension conditions than in either the medium or low evaluation apprehension conditions.

In addition, the effect of the temporal order of the testing of subjects by individual experimenters was assessed. No specific predictions were made concerning this effect.

CHAPTER II

METHOD

Subjects. Two hundred sixteen Ss, from the University of Manitoba, participated in order to partially complete their experimental credits required for the introductory psychology course in which they were enrolled. As no particular prerequisites were stated for participation, Ss of both sexes volunteered. The Ss participated during February and March, 1970.

Experimenters. Eighteen male Es, who were not remunerated, volunteered to participate. These Es were recruited from four different undergraduate psychology courses; two sections of social psychology and two sections of behavioural analysis. Twelve of these Es were from the two behavioural analysis sections.

Design. The 216 Ss were randomly assigned to the 36 cells of a 2 x 3 x 2 x 3 factorial representing two levels of expectancy (low and high), three levels of inducement of bias (low, high, and intentional), two levels of temporal order of testing (early and late), and three levels of set of subjects (low evaluation apprehension, medium evaluation apprehension, and high evaluation apprehension). The 18 Es tested 12 Ss each. Each E tested one S on each of the set, expectancy, and temporal order

conditions. However, he tested Ss in only one of the inducement conditions. Six Es were randomly assigned to each of the inducement conditions.

Task. The task used was a marble-dropping task where Ss dropped marbles, at a high rate of speed, through holes drilled in a table top. Each marble dropped activated a micro-switch which in turn activated an electronic counter. The number of marbles dropped per trial was the unit of measurement. A trial was arbitrarily defined as 60 seconds.

The task used was selected for three reasons:

(1) The "cover story" associated with the marble-dropping task was one which, experience has indicated, was perceived by Es to have a high degree of credulity, and therefore did not arouse suspicion as to the nature of the experiment; (2) The task as presented to the S was fairly ambiguous as to its nature and purpose; and (3) The task was one which had been used in studies (Dusek, 1970; Johnson, 1970) where significant bias effects were obtained.

Inducement of expectancy. The Es' expectancies were induced by verbal instructions given by the principal investigator just prior to the testing of the first S by each E. The principal investigator, who requested each E to appear 30 minutes before the first S was scheduled to appear, explained the task and the procedures to be followed.

In order to produce the expectancy, E was told

that the task was believed to be correlated with intelligence and that the task was one which involved no prior learning, and thus a future intelligence test incorporating such sub-tests as marble-dropping would be less "culturally oriented" than more common present-day tests. Each E was told that the hypothesis was that Ss of higher intelligence would show a greater increase in the rate of marble-dropping across the six trials than would Ss of lower intelligence, the effect reaching its maximum around the sixth trial. It was further explained that during the early part of the twentieth century such researchers as Spearman had attempted to relate intelligence and motor performance, finding low positive correlations. The principal investigator felt that these early researchers had "missed the boat" by failing to take into consideration the changes in response rates over time, and that these changes in rates of responding were not obscured by non-intelligence-related individual differences as were absolute response rates.

The E was told that Ss had been divided into two intelligence groups, split at the median, on the basis of test scores obtained from a "psychometric questionnaire" administered to introductory psychology students at the beginning of the previous term. Each E tested Ss from both of these "intelligence" groups. Each E was given a list containing the name, appointment time, and intelligence group of each S he was to test. Ostensibly this last bit of information was "incidentally" provided, in the

margins of the list rather than as a formal piece of information, for the E's own interest.

Methods of inducement. Three methods of inducing experimenter expectancy were investigated. These methods required different treatments of the Es by the principal investigator.

For Es assigned to the low-inducement condition there was little contact with the principal investigator following the initial inducement. The principal investigator withdrew to an adjacent room (Experimental Room # 2) where he administered questionnaires to Ss who had completed the marble-dropping task in Experimental Room # 3. The investigation was undertaken in four adjacent rooms, all facing the same corridor and all relatively isolated from other research rooms. Ss were tested at approximately 15 minute intervals.

It was assumed that if the principal investigator was convincing in showing great interest in the experiment and in the results, the individual E would become more involved with the running of his Ss and would develop stronger biases than those Es simply given the cover story and left to test their Ss. Thus, for Es assigned to the high-inducement condition, the principal investigator attempted to involve each E by role-playing excitement about the proposal, by communicating to each E the potential importance of the experiment. In addition, the principal investigator entered the main experimental room (# 3) between

the testing of Ss (three times for each E; after the first, third, and sixth S) to ask how things were going, hopefully showing a keen interest in positive results.

For Es assigned to the intentional-bias condition there was a complete briefing prior to the testing of any Ss, and each E was told the true nature of the experiment. The E was informed that the experiment concerned experimenter bias and a brief history of Rosenthal's work was given. The E was told that the principal investigator wished to test whether or not an E could influence his data if he intentionally attempted to manipulate his Ss' responses. The E was asked to attempt to communicate to his Ss the correct type of performance, i. e., increase or no increase in marble-dropping rate. He thus attempted to differentiate his treatment of high and low expectancy Ss, without deviating from the standard instructions, so that high expectancy Ss would increase their performance across trials to a greater extent than low expectancy Ss.

Temporal order of testing. The first six Ss tested by an E were defined as early-tested, and the second six Ss were defined as late-tested. Thus any systematic differences could be assessed and used to provide information as to the question of increasing or decreasing the biasing effects by E across time. Each E tested one S under each of the six possible combinations of evaluation apprehension and expectancy conditions during the testing of his first six Ss. The procedure was then repeated for

the second six Ss. However, the order of the conditions tested during the early testing and during the late testing was random, and subsequently not necessarily the same for an individual E.

Control for recorder and observer bias. Recording of the data was done in a semi-automatic fashion designed to eliminate possible observer and/or recorder errors by the Es. Individual Es did not collect the data, as the counter was placed in experimental room # 4. An additional investigator remained in this room and recorded the number of marbles dropped per trial. This investigator was not aware of any of the conditions so that any recorder or observer errors should have been random in nature across all conditions.

Control for "cheating". So that a later check could be made on such possible "cheating" as overt communication of the hypothesis from E to S, continuous tape recordings were made of all occurrences in room # 3 (unknown to both E and S). For purposes of this study "cheating" was defined as any verbal communication from E to S in which it was stated that the S should increase his speed. As E was instructing S to drop marbles as fast as possible, the manner in which E might have "cheated" would have been for him at any time to have told the S to increase his rate. Although E was cautioned against doing this it is possible that some Es might have told some of their Ss to "speed up". Although the occurrence of cheating was not antici-

pated, it was decided prior to the beginning of the experiment that any comment made by E which directed S to increase or decrease his rate (e. g., "go faster", "pick it up", "come on, you can work faster than that", "try harder", etc.) would be interpreted as "cheating".

Manipulation check on the levels of inducement.

When each E had completed his testing, he was interviewed in a casual manner by the principal investigator. He was queried as to how desirous it had been to him that he obtain positive results. The high-inducement and low-inducement Es were asked whether or not they had intentionally attempted to manipulate their Ss' responses. Each E was also asked about his suspiciousness of the nature of the experiment. Once the interview had been completed the E was debriefed and told the true nature of the experiment.

Sets of Ss. Each S upon reporting for the experiment went to experimental room # 1. He was greeted by an assistant who was kept blind as to the crucial aspects of the experiment. This assistant (a graduate student in English) seated the S and gave him test materials.

The Ss in the low and medium evaluation apprehension conditions completed a questionnaire unrelated to the experimental manipulations. This questionnaire included 52 items asking the S to report his feelings about psychology and about serving in psychological experiments.¹

¹This questionnaire will be discussed in greater detail in Chapter III.

The instructions to this questionnaire explained to the S that he was being asked to complete an attitude scale which was being included in some experiments conducted at the University of Manitoba, as well as at other institutions. The stated purpose in the instructions explained that the experimenters had been asked to include this questionnaire, but that it had no connection with the experiment for which he had "signed up" and which was being conducted in another room. The real purpose for the inclusion of this scale was to provide an unrelated activity of similar length to the test being taken by Ss in the high evaluation apprehension condition.

Following the completion of this questionnaire each S in the low evaluation apprehension condition read an instruction sheet designed to inhibit evaluation apprehension. This communication, patterned after the one used by Rosenberg (1969) for his low evaluation apprehension conditions, led S to believe that the task he was about to engage in was a measure of motor reactions where he was providing data as part of a control or standardization group. These instructions are given in Appendix A.

Following the completion of the same questionnaire, Ss in the medium evaluation apprehension condition read an instruction sheet designed to arouse evaluation apprehension. This communication, patterned after the one used by Rosenberg (1969) for his high evaluation apprehension conditions, led the S to believe that the task he was about

to engage in was correlated with intelligence, and poor performance was related to low intellectual ability. As this study investigated three levels of evaluation apprehension, the third level attempting to arouse even greater apprehension than Rosenberg's "high" manipulation, the labelling of this second level as "medium" is not meant to imply a non-high apprehension arousal. These instructions are given in Appendix B.

The Ss in the high evaluation apprehension condition completed an Otis Self-Administering Test of Mental Ability instead of the questionnaire described above. It is assumed that S recognized this test as an I. Q. scale. Following completion of this test he read the same communication as read by the Ss in the medium evaluation apprehension condition. It was assumed that evaluation apprehension arousal would be even greater for those Ss who were told after the completion of an intelligence test that their performance was again to be evaluated.

Procedure followed by Ss. When Ss had finished the first part of the procedure by completing either the attitude scale or the intelligence scale, they were directed from experimental room # 1 to experimental room # 3. At the completion of the second part (marble dropping in room # 3) all Ss from all conditions reported to room # 2 (the principal investigator's room) where each S completed either the Otis or the attitude scale, whichever test had not already been completed. The administering of the Otis

to all Ss was designed to aid the Es in believing the cover story of a new intelligence test's being developed. It was explained that comparisons were to be made between marble-dropping scores and scores from the "psychometric questionnaires" and between marble-dropping scores and Otis scores. It was further explained that a control for order of testing was being exercised so that some Ss were to receive the Otis prior to the marble-dropping task and some were to receive the Otis following marble dropping. The attitude scale's inclusion was explained as simply being a request of a senior faculty member imposing his wishes on the principal investigator.

Manipulation check on sets of Ss. When each S had completed the last experimental measure, he was asked by the principal investigator to complete a short questionnaire. This questionnaire provided a rough estimate of whether or not the manipulations designed to raise the different levels of evaluation apprehension were successful. This questionnaire is given in Appendix C.

CHAPTER III

RESULTS

The dependent variable was the number of marbles dropped through the holes in the apparatus top during each minute of the task. Of primary interest in all analyses is the change in response rates across trials. As the expectancy inducement for all experimenters explained that the hypothesis predicted greater increases in rates across trials for high than for low expectancy subjects, the presence or absence of expectancy effects is to be determined by observing trials by expectancy interactions. Although the experimenters were told that the principal investigator was concerned only with these across-trials effects, results in line with the induced expectancies should also produce differential absolute response rates for high and low expectancy subjects. Thus, if a subject's response rate increases across trials in line with the experimenter's expectancy, that subject should also show a greater average response rate than shown by a subject not expected to increase his rate across trials. However, the primary test of the presence of expectancy effects concerns only changes or no changes in the dependent variable across trials.

Analyses of variance for repeated measures were computed on the data separately for each method of expect-

ancy inducement; i. e., for low-inducement, high-inducement, and intentional-inducement experimenters. There are two reasons for this approach to the statistical analysis. Firstly, although all subjects were randomly assigned to conditions, and although all experimenters were randomly assigned to the three levels of the inducement condition, experimenters once assigned to an inducement condition could test subjects only within that condition. Thus, for example, while a low-inducement experimenter could test a subject at any level of evaluation apprehension, he could not test a subject assigned to either a high-inducement or an intentional-inducement condition. Myers (1966) suggests that when more than one experimenter is used in a given experiment, subjects should be randomly assigned to both conditions and experimenters. In this particular study, subjects could be randomly assigned to experimenters only within inducement conditions.

Secondly, as the number of experimenters per inducement level was relatively small, six experimenters per level, separate analyses removed the necessity for assuming equivalence of the three groups of experimenters. A total of only 18 experimenters was used, due not only to a finite pool of available volunteers, but as well to the absolute necessity of maintaining the deception concerning the expectancy inducement. For these reasons the means of analysis chosen was seen as providing the least chance of inferential error.

Low-inducement experimenters. The analysis of variance on the dependent measures for those subjects tested by low-inducement experimenters is presented in Table 1.¹ As may be seen, there are potentially three main independent sources of variance: evaluation apprehension, expectancy, and temporal order; and one main correlated source of variance, trials. Only the interaction between evaluation apprehension and expectancy, and the main effect of trials are statistically significant. The mean responses for high, medium, and low evaluation apprehension subjects at high expectancy conditions are: 37.44, 33.33, and 34.24. At low expectancy conditions the respective means are: 33.80, 34.43, and 34.92. That high evaluation apprehension, high expectancy subjects should respond more quickly than other subjects is not specifically related to the experimental hypotheses concerning changes in response rates across trials. The main effect of trials is statistically significant and an inspection of the means indicates that response rates for all subjects increased across trials. This observation is consistent with other studies which have used this task (Parton and Ross, 1965; Stevenson and Hill, 1966).

In Figure 1 are represented the mean responses for high and low expectancy subjects as tested by low-inducement experimenters. The absence of an interaction

¹The mean response rates, pooled over the two levels of temporal order, are presented in Appendix D.

Table 1

Analysis of Variance on Numbers of Marbles Dropped per
Minute by Subjects of Low-Inducement Experimenters

Source	df	SS	MS	F
Evaluation Apprehension	2	221.79	110.90	1.49
Expectancy	1	41.60	41.60	.56
EA EX	2	495.11	247.56	3.32 *
Temporal Order	1	296.71	296.71	3.98
EA TO	2	176.87	88.43	1.19
EX TO	1	166.23	166.23	2.23
EA EX TO	2	160.78	80.40	1.08
Error	60	4474.93	74.58	
Trials	5	365.86	73.17	20.07 **
EA TR	10	35.62	3.57	.98
EX TR	5	8.13	1.63	.47
EA EX TR	10	25.91	2.59	.71
TO TR	5	12.41	2.48	.68
EA TO TR	10	33.76	3.37	.92
EX TO TR	5	5.54	1.10	.30
EA EX TO TR	10	20.70	2.07	.57
Error	300	1093.73	3.65	

* $p < .05$

** $p < .001$

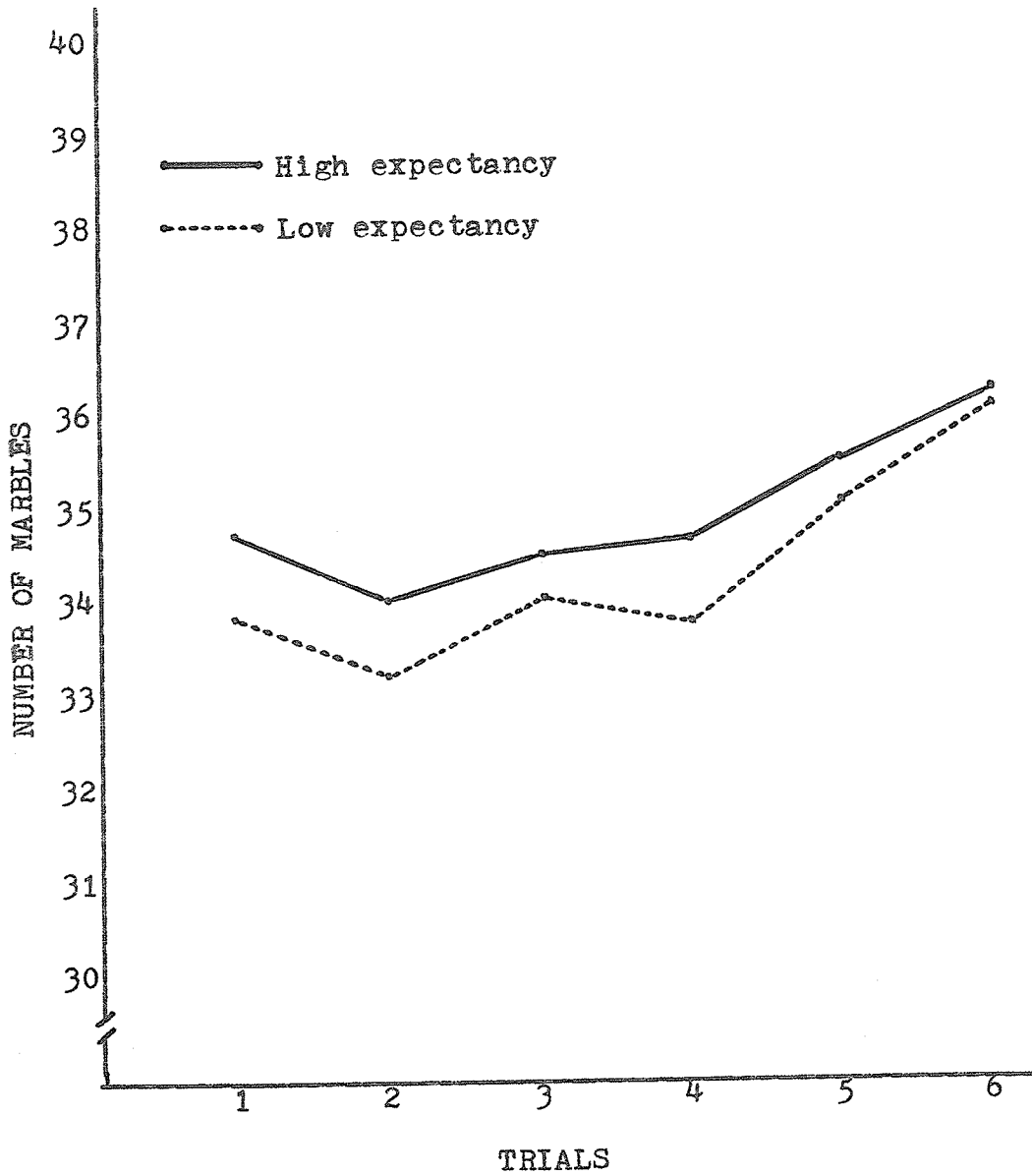


Figure 1. Mean responses for subjects tested by low-inducement experimenters.

between expectancy and trials fails to support the hypothesis of an expectancy effect. It should be noted, however, that the mean responses for high expectancy subjects increased in magnitude to a lesser extent than did the mean responses for low expectancy subjects. This slight difference, in the opposite direction from the hypothesis, is not statistically significant.

High-inducement experimenters. In Table 2 is presented the analysis of variance on the dependent measures for those subjects tested by high-inducement experimenters.¹ As before, there are the same three main independent and one main correlated sources of variance. In this analysis the main effects of evaluation apprehension and trials as well as the interaction among evaluation apprehension, expectancy, and trials are statistically significant. As in the analysis of the data of low-inducement experimenters it was observed that response rates for all subjects increased across trials. The mean response rates for high, medium, and low evaluation apprehension subjects (36.17, 33.29, 34.92) are significantly different. While the experimental hypotheses were not concerned with response rates other than across trials, an interaction between expectancy and evaluation apprehension, as was observed in the analysis of data obtained by low-inducement experimenters, would not be discrepant with the hypotheses. However, the finding

¹The mean response rates, pooled over the two levels of temporal order, are presented in Appendix E.

Table 2

Analysis of Variance on Numbers of Marbles Dropped per
Minute by Subjects of High-Inducement Experimenters

Source	df	SS	MS	F
Evaluation Apprehension	2	604.54	302.27	3.37 *
Expectancy	1	107.02	107.02	1.19
EA EX	2	347.89	173.95	1.94
Temporal Order	1	204.20	204.20	2.28
EA TO	2	60.48	30.24	.34
EX TO	1	.00	.00	.00
EA EX TO	2	70.24	35.12	.39
Error	60	5378.46	89.64	
Trials	5	244.33	48.87	10.81 ***
EA TR	10	31.86	3.19	.71
EX TR	5	9.47	1.89	.42
EA EX TR	10	105.65	10.57	2.34 **
TO TR	5	1.79	.36	.08
EA TO TR	10	28.07	2.81	.62
EX TO TR	5	22.70	4.54	1.00
EA EX TO TR	10	41.93	4.19	.93
Error	300	1356.04	4.52	

* $p < .05$

** $p < .02$

*** $p < .001$

in this analysis of a significant main effect of evaluation apprehension is somewhat surprising. Rosenberg's (1969) interpretation of the concept of evaluation apprehension would not seem to be consistent with the observation that the medium evaluation apprehension subjects responded at the slowest rate.

The significant interaction among evaluation apprehension, expectancy, and trials supports the hypothesis that more positive expectancy effects occur with increasing degrees of evaluation apprehension. In Figures 2A, 2B, and 2C are represented the mean responses for high and low expectancy subjects as tested by high-inducement experimenters. It may be observed that high evaluation apprehension subjects under high expectancies increased response rates across trials to a greater extent than did high evaluation apprehension subjects under low expectancies. For medium evaluation apprehension subjects this difference is not as great, and it should be noted that the increase for the high expectancy subjects did not occur until the sixth trial. However, for low evaluation apprehension subjects the direction of these means is reversed.

Intentional-inducement experimenters. The analysis of variance performed on the dependent measures for those subjects tested by intentional-inducement experimenters is presented in Table 3.¹ The main effect of trials

¹The mean response rates, pooled over the three levels of evaluation apprehension, are presented in Appendix F.

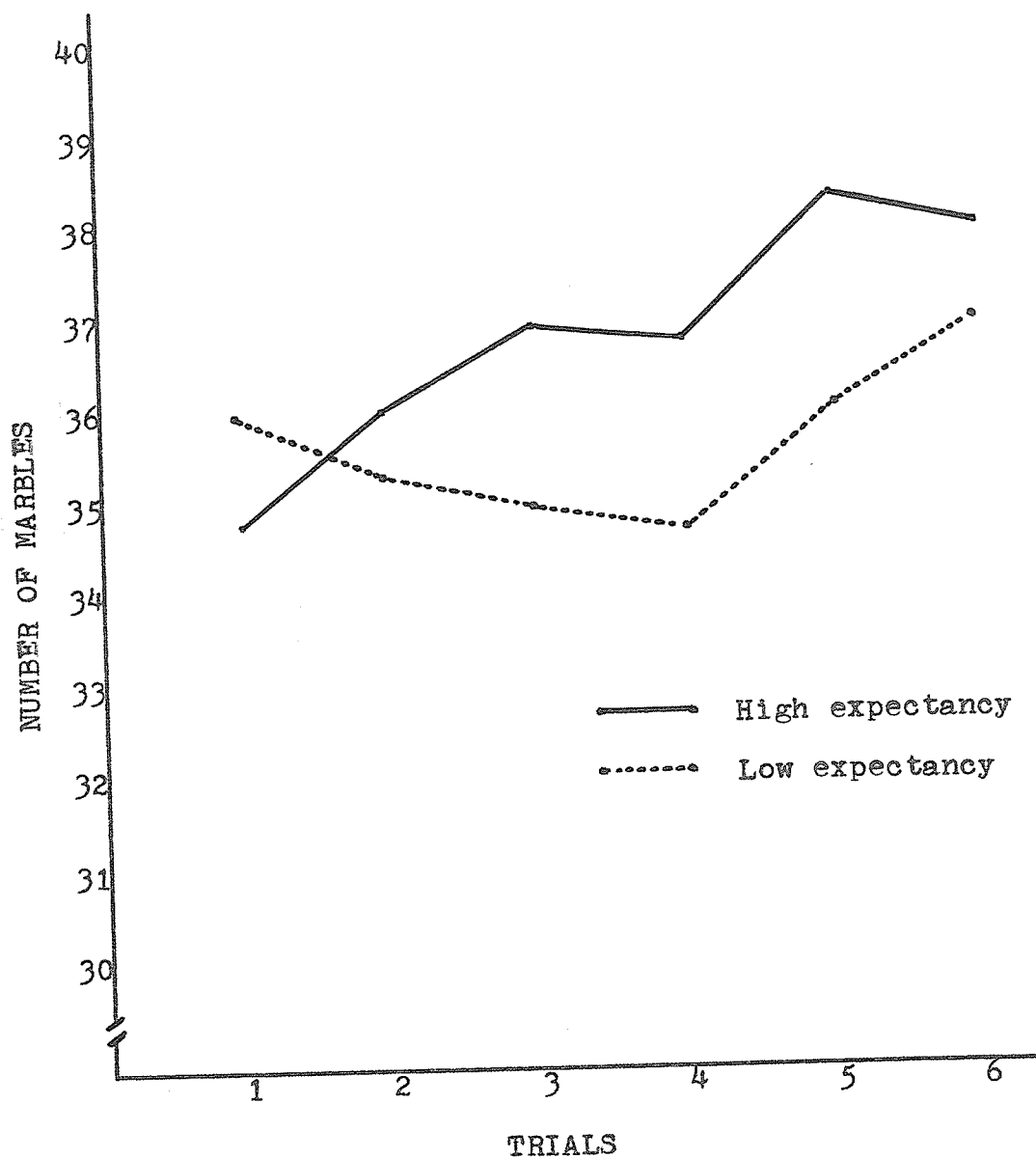


Figure 2A. Mean responses for high evaluation apprehension subjects tested by high-inducement experimenters.

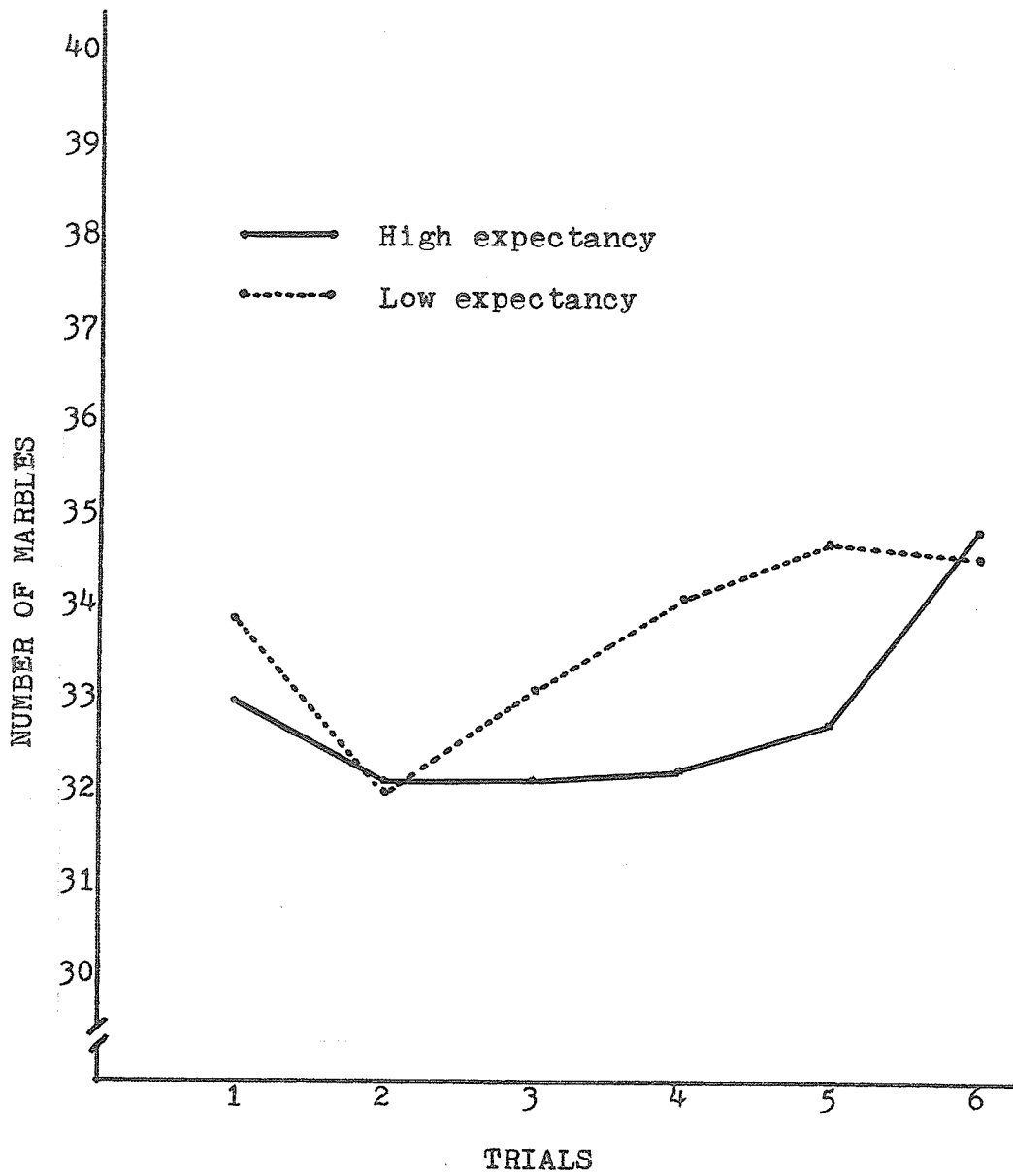


Figure 2B. Mean responses for medium evaluation apprehension subjects tested by high-inducement experimenters.

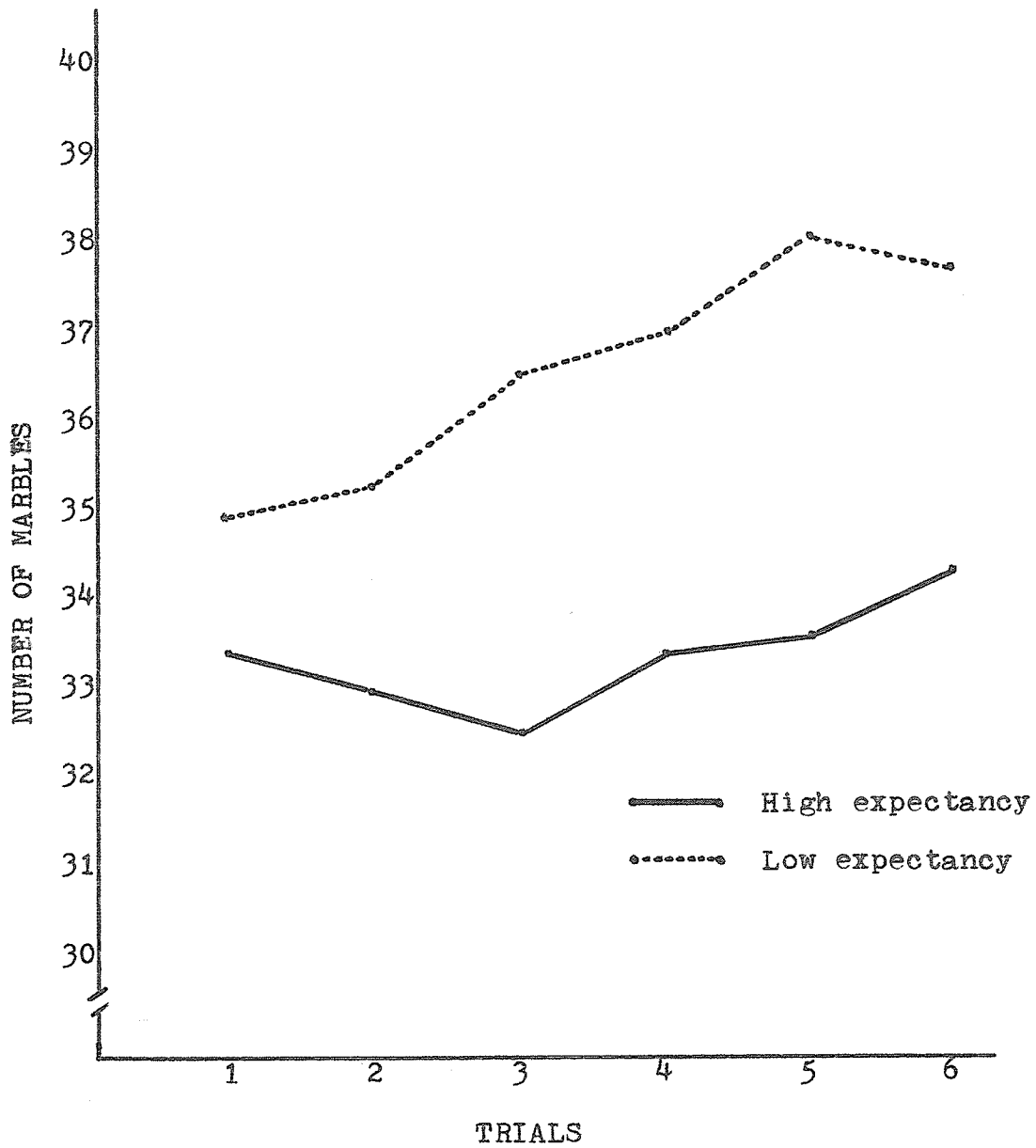


Figure 2C. Mean responses for low evaluation apprehension subjects tested by high-inducement experimenters.

Table 3

Analysis of Variance on Numbers of Marbles Dropped per Minute by Subjects of Intentional-Inducement Experimenters

Source	df	SS	MS	F
Evaluation Apprehension	2	179.68	89.84	1.07
Expectancy	1	46.02	46.02	.55
EA EX	2	150.50	75.25	.90
Temporal Order	1	132.23	132.23	1.57
EA TO	2	37.85	18.93	.23
EX TO	1	819.50	819.50	9.74 **
EA EX TO	2	1.69	.84	.01
Error	60	5047.52	84.13	
Trials	5	507.49	101.49	27.47 ***
EA TR	10	47.46	4.74	1.28
EX TR	5	33.13	6.63	1.79 *
EA EX TR	10	25.96	2.60	.70
TO TR	5	10.15	2.03	.55
EA TO TR	10	37.89	3.79	1.03
EX TO TR	5	4.37	.87	.24
EA EX TO TR	10	25.41	2.54	.69
Error	300	1108.65	3.70	

* $p < .12$

** $p < .005$

*** $p < .001$

and the interaction between expectancy and temporal order are statistically significant. The interaction between expectancy and trials approaches significance ($p < .12$). As was the case for subjects tested by low-inducement and by high-inducement experimenters response rates for all subjects increased across trials. The significant interaction between expectancy and temporal order is not specifically related to the experimental hypotheses of changes in response rates across trials. The mean responses for high expectancy subjects tested early and tested late by the experimenters are 34.90 and 33.25. The respective means for low expectancy subjects are 31.49 and 35.35.

The mean responses for high and low expectancy subjects as tested by intentional-inducement experimenters are represented in Figure 3. It may be observed that low expectancy subjects increased their rates to a greater extent than did high expectancy subjects. This finding is directly opposite to the experimental hypothesis that high expectancy subjects would increase their responses to a greater extent than low expectancy subjects. While this interaction only approaches significance ($p < .12$) in the analysis of variance, an inspection of Figure 3 indicates a considerable change in the response rate from the first to the sixth trial for high expectancy subjects relative to the change for low expectancy subjects. A t-test performed on the difference scores for high and low expectancy subjects indicates a significant difference

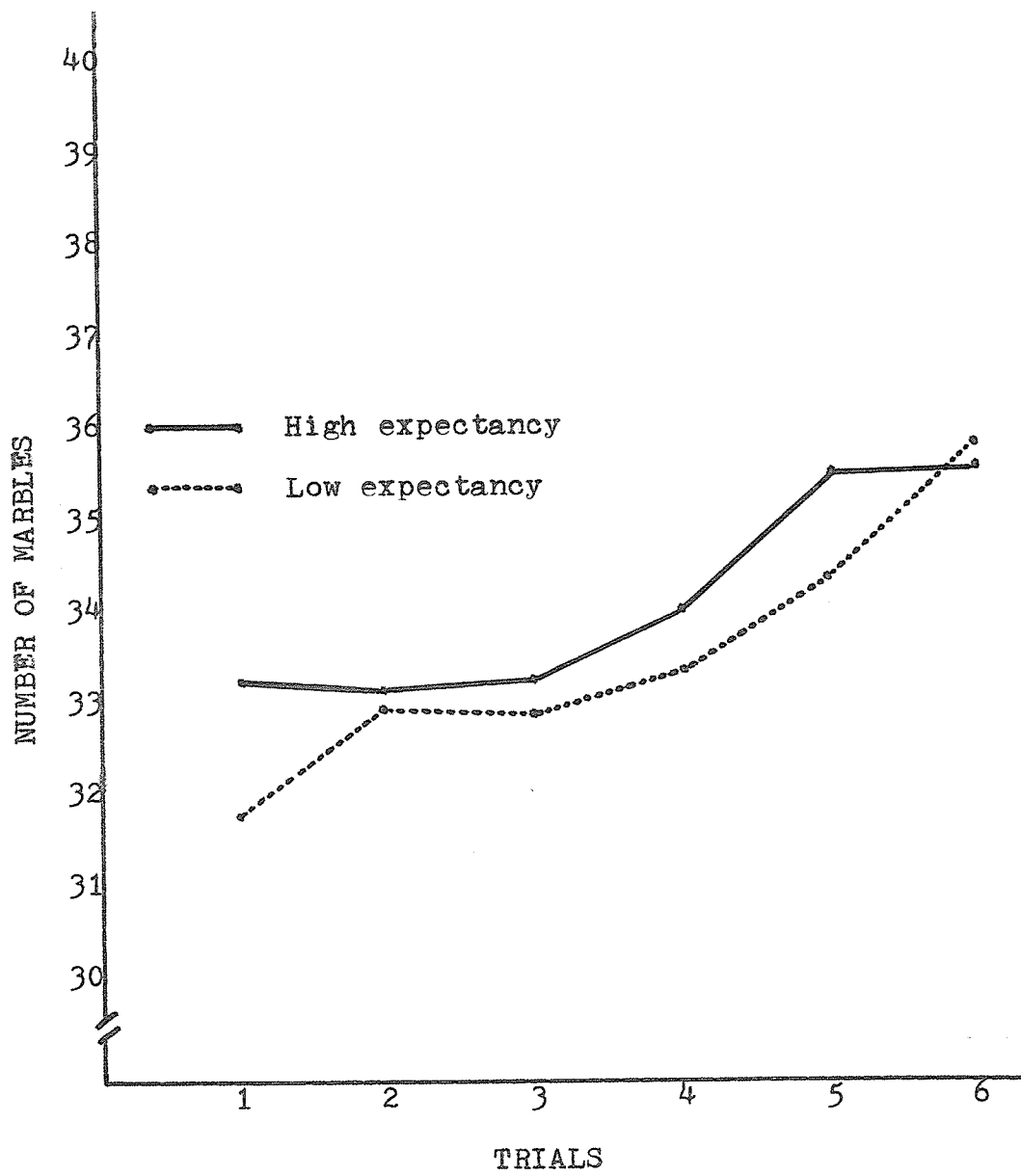


Figure 3. Mean responses for subjects tested by intentional-inducement experimenters.

between the two groups ($t = 2.38$, $df = 70$, $p < .02$, two-tailed). These difference scores were computed for each subject by subtracting the number of responses for minute one from the number for minute six.

Post-experimental interviews of experimenters and analysis of tape recordings. All experimenters were interviewed by the principal investigator following the experimental sessions. None of the low-inducement nor high-inducement experimenters reported that they had attempted to intentionally manipulate any of their subjects' responses. Two experimenters, one low-inducement and one high-inducement, reported suspiciousness as to the nature of the experiment. In both cases the experimenters reported that they were suspicious, prior to the experimental session, that they themselves might be the real subjects of the experiment. In neither case was the experimenter certain until he was debriefed. The data obtained for these experimenters were not discarded from any of the analyses.

While no objective evidence can be cited to support the observation, it is the principal investigator's impression, following the interviewing, that the differential inducements were successful. The high-inducement experimenters appeared to have been more interested in the experiment and more concerned as to its outcome than were the low-inducement experimenters.

All of the experimental sessions were tape-recorded without the awareness of either experimenters or

subjects. An analysis of these recordings revealed that only once did an experimenter verbalize any comment to a subject which approached the criterion for "cheating" as outlined in the previous chapter. One of the intentional-inducement experimenters made the comment "keep 'em comin' boy" to one of his subjects. He did not repeat the comment nor make any similar statement to that particular subject nor to any of the other subjects he tested. The data for this subject were not discarded from any of the analyses.

The analysis of the recordings revealed two un-predicted but interesting occurrences. Firstly, it was observed that two of the six low-inducement experimenters asked the names of some of their subjects (three instances for one experimenter, four for the other) only after the subjects had been tested. It cannot be determined whether these experimenters assumed the subjects in question were the next scheduled subjects, or whether these experimenters were not concerned about the "intelligence" group of these subjects.

Secondly, it was observed that the intentional-inducement experimenters emitted more bits of verbalizations than did the other experimenters. A count was taken of the number of times each experimenter made any sort of comment other than to respond to a subject's unsolicited comment. The mean number of comments made by the intentional-inducement experimenters was 49.17 or 3.10 comments per subject. The high-inducement experimenters commented

a mean number of times of 10.83 or .90 times per subject, and the low-inducement experimenters 8.67 or .74 times per subject. Taking the number of comments made by each experimenter as the measure, t-tests reveal differences between intentional-inducement experimenters and high-inducement experimenters ($t = 2.04$, $p < .08$, two-tailed), between intentional-inducement experimenters and low-inducement experimenters ($t = 2.05$, $p < .08$, two-tailed), and between intentional-inducement experimenters and high- and low-inducement experimenters combined ($t = 2.89$, $p < .01$, two-tailed). No differences were observed between high- and low-inducement experimenters nor within experimenters when testing high and low expectancy subjects.

Post-experimental questionnaire for subjects.

The subjects were first asked: "How concerned were you with your performance on the Marble Test?" Subjects could check one of four categories, where "0" was "not concerned", "+1 neither concerned nor not concerned", "+2 somewhat concerned", and "+3 substantially concerned". The median response was 2.44 indicating considerable concern on the part of the subjects as to their performances. However, a Kruskal-Wallis one-way analysis of variance corrected for ties (Siegal, 1956), performed on the responses checked by high, medium, and low evaluation apprehension subjects, yielded no significant differences among the three groups.

The second question asked the subjects was: "How did the written instructions explaining the purpose

of the experiment affect you as you were about to take the Marble Test?" Responses could vary on a five-point scale from "-2 Aroused my anxiety to a great extent" to "+2 Greatly reassured me". The respective medians for the high, medium, and low evaluation apprehension subjects were: $-.07$, -1.44 , and $+.02$. These medians are significantly different, as tested by the Kruskal-Wallis one-way analysis of variance corrected for ties ($H = 8.19$, $p < .02$).

The third question asked: "Did the experimenter do anything to help you perform better on the Marble Test?" A total of 72 subjects responded "yes". A Chi-Square yielded no significant differences ($\chi^2 = 3.06$) between those subjects tested by the three groups of differentially-induced experimenters. Those subjects who elaborated on this item tended to respond with comments such as "Yes, he was polite" or "Yes, he carefully explained the instructions".

The fourth question asked: "Did you look to the experimenter for clues as to how you might best perform on the Marble Test?" Only 23 subjects responded "yes" to this item and there were no differences among the high, medium, and low evaluation apprehension subjects ($\chi^2 = 2.32$). Few subjects elaborated on their responses.

The final question asked: "What do you think was the purpose of this experiment?" In almost all cases the subjects responded with answers congruent with what they had been instructionally led to believe. Only three subjects indicated awareness or partial awareness. These

three comments were: (1) "To measure how hard people try on experiments", (2) "I think the purpose of the experiment was to see how the experimenters actions & what he said effected [sic] me.", and (3) "It was probably to check the student's ability to perform after he has been told that if he does poorly he is lacking in intelligence--because it concerns him & his worth at the U. (supposedly) it may affect performance (also to see exactly how much students believe in these tests)." That only three subjects admitted awareness should not necessarily be interpreted as indicating complete effectiveness of the deception manipulations. Golding and Lichtenstein (1970) have recently reported low rates of "confessed awareness-suspicion" in an experiment where subjects had been "tipped off" prior to the deception manipulations. Their post-experimental measures were of a more sophisticated type than those used in the present study.

Subjects' attitudes toward psychology. The attitude scale completed prior to the experimental manipulations by the low and medium evaluation apprehension subjects and following the manipulations by the high evaluation apprehension subjects was a 52-item scale (Adair and Fenton, 1970) developed to measure attitudes toward psychology and psychological research. Adair (1970) observed relationships between this scale and the degree of awareness of the experimenter's hypothesis as well as willingness to co-operate with that hypothesis. However, not

all of the relationships he observed were consistent as to direction. For example, more positive attitudes were related to greater conformity in a conformity study, but to less conditioning with a standard verbal conditioning task.

Pearson product-moment correlation coefficients were computed for the attitude scores and the difference scores for high and low expectancy subjects tested by the three differentially-induced groups of experimenters. Two significant relationships were observed. With high-inducement experimenters, attitude scores for the low expectancy subjects were inversely related to increases in response rates ($r = -.58$, $N = 36$, $p < .01$). Thus it would appear that the more positive a subject's attitudes toward psychology, the more consistent was his performance on the marble task, where the experimenter's hypothesis was that he would not increase his rate of responding. However, with low-inducement experimenters, the relationship between attitude scores and performance for low expectancy subjects was reversed ($r = +.37$, $N = 36$, $p < .05$). It should be remembered that no significant expectancy effect was observed for low-inducement experimenters, but that the means were in the direction opposite to the experimenters' hypotheses.

Interpretation of these observations is made somewhat difficult by the absence of relationships in other conditions, in particular the absence of relationships between attitude scores and performance scores for high expectancy subjects tested by either low- or high-induce-

ment experimenters. Perhaps the most parsimonious explanation is that the two observed relationships were spurious.

CHAPTER IV

DISCUSSION

Two rough checks on the effectiveness of the experimental manipulations were undertaken while the data were being collected: the experimenters were interviewed following their testing of subjects, and the subjects completed a short post-experimental questionnaire. While the interviews indicated that the differential inducements of the expectancies had been successful, the analyses of responses to the questionnaire raised some doubt as to the effectiveness of the evaluation apprehension manipulations.

The question which asked subjects to report their degree of concern about their performance was not differentially responded to by the high, medium, and low evaluation apprehension subjects. Holmes and Appelbaum (1970) recently reported that groups of subjects with different prior experimental experiences, and whose performances were significantly differentially affected, did not report differences in effort on a post-experimental questionnaire. They interpret this observation in terms of the strong demand characteristics of the post-experimental situation which require subjects not to report any reduction in effort. In the present investigation an alternative explanation is that the differential evaluation apprehension manipu-

lations were not successful.

That the medium evaluation apprehension subjects should have indicated greater anxiety arousal over the written instructions than did the other two groups of subjects is somewhat surprising. However, it may have been due to the fact that the high evaluation apprehension group read the instructions immediately following their taking an intelligence scale and were thus less surprised than the mediums, whose anxiety had not been previously aroused by their having been tested. It is a surprising observation, however, that the high evaluation apprehension subjects did not report greater anxiety than did the low evaluation apprehension subjects. Again, it may have been due to the manipulation where the apprehension of the subjects may have already been aroused by the Otis prior to their reading of the evaluation instructions. In this case the subjects may have reported accurately that the instructions did not raise anxiety. Nonetheless the responses to this second question on the post-experimental questionnaire suggest a second alternative explanation which cannot be ruled out by the present data. It is possible that the group of subjects labelled "medium" actually had their evaluation apprehension aroused to the greatest extent of the three groups of subjects.

A number of interesting and important observations are to be made from the analyses on the rates of marble dropping. The effects of experimenter expectancy were

measured in a number of experimental conditions. In some of these conditions positive expectancy effects were observed, in some conditions effects opposite to the experimenters' expectancies occurred, and in some conditions the induced expectancies seemed to have no effect at all. The patterns of the occurrence and the direction of these experimenter expectancy effects shed considerable light on the phenomenon, and as well raise a number of questions about the phenomenon which require empirical answers.

In the first hypothesis greater improvement in the rate of marble-dropping across trials for subjects in high expectancy conditions than for subjects in low expectancy conditions was predicted. The only conditions which showed expectancy effects in the direction of the experimenters' expectancies were those conditions involving high-inducement experimenters testing medium and high evaluation apprehension subjects. Even here it would appear that expectancy effects with medium evaluation apprehension subjects were extremely weak, if not spurious. For low evaluation apprehension subjects tested by high-inducement experimenters and for all subjects tested by low-inducement and by intentional-inducement experimenters there were no positive expectancy effects. Although not all of these latter conditions showed significant differences across trials between high and low expectancy subjects, the means in all cases were in directions opposite to the induced expectancies. It would appear from the weak support for

this hypothesis that experimenter expectancy, as an experimental phenomenon, may be effective in only certain types of experimental situations. This should not necessarily be interpreted as suggestive of a weak phenomenon occurring only infrequently in general psychological research. To the extent that the observations of this investigation can be generalized beyond the present data it seems crucial to consider first under what conditions the phenomenon was a significant source of variance.

In the analysis performed on the responses of those subjects tested by low-inducement experimenters, the absence of any experimental effects across trials, as noted above, fails to support the hypothesis of overall expectancy effects. There was, however, a significant interaction between expectancy and evaluation apprehension. This finding would suggest that the induced expectancy had some effect on the obtained results, but that the effect was not strong enough to cause the subjects to respond directly in line with that expectancy. That is to say, the induced expectancy may have affected the overall rates of response but did not effect appropriate increases or decreases in rates of responding across trials.

In the analysis performed on the responses of those subjects tested by high-inducement experimenters, the significant interaction of evaluation apprehension, expectancy, and trials was in the predicted direction. This finding supports both the hypothesis of expectancy

effects for high-inducement experimenters and the hypothesized interaction between evaluation apprehension and expectancy across trials. The results of this analysis closely approximate the results obtained by Rosenberg (1969), who used the person-perception task, and therefore add weight to his contention that evaluation apprehension is a necessary mediator of experimenter expectancy.

In hypothesis 2A were predicted less positive expectancy effects for low-inducement conditions than for high-inducement conditions. The method of analysis, performed separately on the sets of data collected by differentially-induced experimenters, did not allow a statistical test of this predicted interaction. However, the means across trials were opposite in direction from the experimenters' expectancies for high and low expectancy subjects tested by low-inducement experimenters. This obvious failure of the induced expectancy to effect results in line with the experimenters' hypotheses can only be interpreted as indicating no positive expectancy effects for low-inducement experimenters. In hypothesis 2B it was predicted that the expectancy effect would be most positive where experimenters intentionally attempted to bias their results. While a statistical test of the interaction between methods of inducement and expectancy was not made, it is obvious from the significantly reversed expectancy effect for the intentional-inducement conditions that this prediction was not correct.

While hypotheses 3A and 3B, where increasingly positive expectancy effects for increasing evaluation apprehension arousal were predicted, found support from the analysis of the responses of subjects tested by high-inducement experimenters, there was no interaction across trials between expectancy and evaluation apprehension for subjects tested by low-inducement experimenters. The absence of this interaction may perhaps have been due to a failure on the part of the experimenters to transmit cues to the subjects. If in fact the experimenters did not behave differentially following the inducement, it would not have been possible for any of the subjects, no matter under what level of evaluation apprehension, to respond in line with the experimental hypothesis. This interpretation, however, is restricted by the observed significant interaction between evaluation apprehension and expectancy. While there were no differences across trials, this two-way interaction indicates that high evaluation apprehension, high expectancy subjects dropped more marbles for all trials than did subjects in the other five conditions. One way to interpret this finding is to assume that the low-inducement experimenters did in fact behave somewhat differentially for high and low expectancy subjects, but that the cues thus transmitted were either so few or so ambiguous that they were picked up only by high evaluation apprehension subjects, and even then were not interpreted "correctly".

On the other hand, if this latter interpretation

is correct one would expect that these transmitted cues should also have been at least partially picked up by the medium evaluation apprehension subjects. This did not appear to have happened. It may have been that the medium evaluation apprehension manipulation was generally ineffective in arousing anxiety over performance. The results for medium evaluation apprehension subjects tested by high-inducement experimenters were as well somewhat ambiguous. It seems quite possible that the instructions telling subjects that their intelligence was to be measured simply were not believed by some portion of the subjects in the medium condition. Rosenberg (1969) reported suspicion with a highly similar communication. It will be remembered that the rationale presented, in the first chapter of this manuscript, for the high evaluation apprehension manipulation concerned the credulity to the subjects of Rosenberg's communication. In the high evaluation apprehension condition subjects actually took an intelligence test which presumably added plausibility to the instructions.

However, that the medium evaluation apprehension communication was effective is suggested by the results of one of the questions on the post-experimental questionnaire which indicated that the mediums reported more anxiety over the instructions than did subjects in the other two groups. It is the opinion of the author, however, that if this group were, in fact, suspicious over the instructions, the subjects would have been more rather than less inclined

to pick out the demand characteristics of the question which required that they report that they were made anxious by the instructions. Nonetheless, an alternative explanation to the suspicion hypothesis is that the mediums were so aroused that their performance was inhibited. This second alternative explanation cannot be ruled out by the present data.

Looking at the results of the analyses on responses made by subjects tested by low- and high-inducement experimenters, where no across-trials expectancy effects were observed with low-inducement experimenters and where a significant interaction among expectancy, evaluation apprehension, and trials was observed with high-inducement experimenters, a tentative interpretation can be made with reference to two perhaps-necessary variables for the mediation of experimenter expectancy. It may be that experimenter expectancy can be mediated only when both the experimenter and the subject are concerned with the outcome of the data-collection process. Neither having an experimenter who is concerned that the data support his hypothesis, nor having a subject who is concerned with his own performance may alone be sufficient. In most reported studies which have shown expectancy effects neither evaluation apprehension nor some other subject set was directly manipulated. Most commonly a subject completes the expectancy task under somewhat ambiguous circumstances and with few preliminary instructions. It may well be that a subject is set to

look for cues as to how to perform under these conditions. In a somewhat unstructured situation the subject may look to the experimenter and conform to his expectancies. In the present investigation the medium and high evaluation apprehension subjects were led to believe that they were to be evaluated and consequently they should have looked to the experimenter for cues as to how best to perform. If, however, the mediums were suspicious that the instructions were included to test whether their performance would be subsequently affected, they may have resisted conforming to the experimenters' cues.

Similarly, it may well have been the case in most experiments where expectancy has been demonstrated that the expectancy was induced under conditions approximating the high-inducement manipulations of this investigation. It is possible that those principal investigators who have found expectancy effects may have shown considerable enthusiasm when inducing the expectancies in their experimenters. Earlier in this manuscript it was suggested that the expectancy effect may increase as set to transmit and set to receive information about performance are incremented. The present investigation indicates that certain modifications need be made in this formulation. It would appear that unless some minimum of both sets is achieved experimenter expectancy is not mediated. An experimental test of this contention would seem to be extremely important at this point.

One of the most interesting observations to be made from the present data comes from the analysis performed on the responses made by those subjects tested by intentional-inducement experimenters. Before attempting to interpret this finding a caution should be made concerning this observation. As the effect has an associated probability of only 0.12, although a t-test performed on the difference scores for minutes six minus one was highly significant, any interpretation must be made, awaiting replication of the observation, with a low level of confidence.

To the extent that this observation is reliable, however, the validity of Barber and Silver's (1968a; 1968b) criticism, concerning intentionality, of the expectancy literature is in question. Barber and Silver suggested that most expectancy effects may have been due to the experimenters' intentionally attempting to manipulate subjects' responses. The present study indicates that for the marble-dropping task, experimenters may in fact have not been able to intentionally manipulate their results, at least not in the direction of their hypotheses.

Why the experimenters failed to produce results in line with their conscious attempts, and why the results tended to be in a direction opposite to those attempts has interest beyond the above-mentioned criticism of the literature made by Barber and Silver. Orne (1962) suggests that subjects may frequently bend over backwards in an attempt to be "honest" when they have been made aware of

the experimenter's hypothesis because of blatantly obvious demand characteristics. It may have been in the present study that subjects reacted against conspicuous cues being emitted by intentional-inducement experimenters. This interpretation is strongly supported by the observation that intentional experimenters were much more verbal than were other experimenters. One can only speculate, of course, as to the internal state of the subjects at the time of testing. However, it may have been that subjects perceived the experimenter as not fulfilling his role as a scientist, and consequently reacted against his attempts at data manipulation. While these experimenters talked as much to the high expectancy subjects as to the low expectancy subjects, the total pattern of cues may have been such as to cause the low expectancy subjects to have perceived the experimenters' attempts at cue passing as deliberate distractions, aimed at assessing their abilities to concentrate on the task at hand.

It should be noted that in addition to the reversed expectancy effects across trials an interaction was observed between expectancy and temporal order of testing. For the first six subjects tested by each experimenter the high expectancy subjects responded at a faster rate than the low expectancy subjects. For the subjects tested later by each experimenter this finding was reversed. This is the only finding in all of the analyses performed that indicates differences due to temporal order. Perhaps the

intentional-inducement experimenters realized early that they were not getting the desired effect and changed the pattern of cues they were emitting.

If the data elicited by intentional-inducement experimenters is reliable, it would appear that the assumption of a monotonic relationship between the degree to which the experimenter is committed to the substantiation of the hypothesized relationship and the degree of the expectancy effect, is in error. It may well be that beyond some to-be-determined point greater concern on the part of the experimenter toward his subject's performance may lead to results less in line with his expectancies.

It would appear that mediation of experimenter expectancy is a complex process and that a model assessing the similarity between the expectancy and the results may not be the best one with which to investigate the expectancy phenomenon. For example, it is possible that one experimenter may bias his results in the opposite direction to the results effected by another experimenter. In this case it would be in error to conclude that experimenter expectancy had no effect on the data-collection process.

The unpredicted observation that the intentional-inducement experimenters made more talking behaviours than the other experimenters suggests that it may have been these behaviours that caused the subjects not to respond in the directions of the expectancies. Subjects may react against any verbalizations emitted by the experimenter during

the completion of the experimental task. This could be because of a subject's perception of the role of an experimenter which does not include this type of behaviour.

Thus, the question should be raised as to what results would have been obtained had these experimenters been cautioned against such verbalizations. This, of course, is an empirical question that should be answered.

References

- Adair, J. G. Pre-experiment attitudes toward psychology as a determinant of experimental results: verbal conditioning of aware subjects. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, in press.
- Adair, J. G., & Epstein, J. S. Verbal cues in the mediation of experimenter bias. Psychological Reports, 1968, 22, 1045-1053.
- Adair, J. G., & Fenton, D. P. Subject's attitudes toward psychology as a determinant of experimental results. Paper delivered at the Midwestern Psychological Association meetings, Cincinnati, 1970.
- Adler, N. E. The influence of experimenter set and subject set on the experimenter expectancy effect. Unpublished AB thesis, Wesleyan College, 1968.
- Barber, T. X., Caverly, D. S., Forgione, A., McPeake, J. D., Chaves, J. F., & Brown, B. Five attempts to replicate the experimenter bias effect. Journal of Consulting and Clinical Psychology, 1969, 33, 1-6.
- Barber, T. X., & Silver, M. J. Fact, fiction, and the experimenter bias effect. Psychological Bulletin

- Monograph Supplement, 1968, 70, 1-29. (a)
- Barber, T. X., & Silver, M. J. Pitfalls in data analysis and interpretation: a reply to Rosenthal. Psychological Bulletin Monograph Supplement, 1968, 70, 48-62. (b)
- Dusek, J. B. Experimenter bias in performance of children at a simple motor task. Developmental Psychology, 1970, in press.
- Friedman, N. The social nature of psychological research: The psychological experiment as a social interaction. New York: Basic Books, 1967.
- Golding, S. L., & Lichtenstein, E. Confession of awareness and prior knowledge of deception as a function of interview set and approval motivation. Journal of Personality and Social Psychology, 1970, 14, 213-223.
- Holmes, D. S., & Appelbaum, A. S. Nature of prior experimental experience as a determinant of performance in a subsequent experiment. Journal of Personality and Social Psychology, 1970, 14, 195-202.
- Johnson, R. W. Subject response as affected by experimenter expectancy, sex of experimenter, and verbal reinforcement. Canadian Journal of Behavioural Science, 1970, 2, 60-66.
- Johnson, R. W., & Adair, J. G. The effects of systematic recording error vs. experimenter bias on latency of word association. Journal of Experimental

Research in Personality, 1970, in press.

- Levy, L. H. Reflections on replications and the experimenter bias effect. Journal of Consulting and Clinical Psychology, 1969, 33, 15-17.
- Marwit, S. J., & Marcia, J. E. Tester bias and response to projective instruments. Journal of Consulting Psychology, 1967, 31, 253-258.
- Masling, J. Differential indoctrination of examiners and Rorschach responses. Journal of Consulting Psychology, 1965, 29, 198-201.
- Myers, J. L. Fundamentals of experimental design. Boston: Allyn and Bacon, 1966.
- Orne, M. On the social psychology of the psychological experiment with particular reference to demand characteristics and their implications. American Psychologist, 1962, 11, 776-783.
- Parton, D. A., & Ross, A. O. Social reinforcement of children's motor behavior: a review. Psychological Bulletin, 1965, 64, 65-73.
- Raffetto, A. M. Experimenter effects on subjects' reported hallucinatory experiences under visual and auditory deprivation. Paper delivered at the Midwestern Psychological Association meetings, Chicago, 1968.
- Rosenberg, M. J. When dissonance fails: on eliminating evaluation apprehension from attitude measurement. Journal of Personality and Social Psychology, 1965, 1, 28-42.

- Rosenberg, M. J. The conditions and consequences of evaluation apprehension. In R. Rosenthal and R. L. Rosnow (Eds.) Artifact in behavioral research. New York: Academic Press, 1969, 279-349.
- Rosenthal, R. On the social psychology of the psychological experiment: The experimenter's hypothesis as unintended determinant of experimental results. American Scientist, 1963, 51, 268-283.
- Rosenthal, R. The effect of the experimenter on the results of psychological research. Bulletin of the Maritime Psychological Association, 1964, 13, 1-39.
- Rosenthal, R. Experimenter effects in behavioral research. New York: Appleton-Century-Crofts, 1966.
- Rosenthal, R. Covert communication and tacit understanding in the psychological experiment. Psychological Bulletin, 1967, 67, 356-367.
- Rosenthal, R. Experimenter expectancy and the reassuring nature of the null hypothesis decision procedure. Psychological Bulletin Monograph Supplement, 1968, 70, 30-47.
- Rosenthal, R. Interpersonal expectations: effects of the experimenter's hypothesis. In R. Rosenthal and R. Rosnow (Eds.) Artifact in behavioral research. New York: Academic Press, 1969, 182-277.
- Rosenthal, R., & Fode, K. L. (Psychology of the scientist:

V) Three experiments in experimenter bias.

Psychological Reports, 1963, 12, 491-511. (a)

Rosenthal, R., & Fode, K. L. The effect of experimenter bias on the performance of the albino rat.

Behavioral Science, 1963, 8, 183-189. (b)

Rosenthal, R., & Jacobson, L. Self-fulfilling prophecies in the classroom: teacher's expectations as unintended determinants of pupils' intellectual competence. In M. Deutsch, I. Katz, and A. R. Jensen (Eds.) Social Class, Race, and Psychological Development. New York: Holt, Rinehart and Winston, 1968.

Rosenthal, R., & Lawson, R. A longitudinal study of the effects of experimenter bias on the operant learning of laboratory rats. Journal of Psychiatric Research, 1964, 2, 61-72.

Rosenthal, R., Persinger, G. W., Mulry, R. C., Vikan-Kline, L., & Grothe, M. Emphasis on experimental procedure, sex of subjects, and the biasing effects of experimental hypotheses. Journal of Projective Techniques and Personality Assessment, 1964, 28, 470-473.

Rosenthal, R., Persinger, G. W., Vikan-Kline, L., & Mulry, R. C. The role of the research assistant in the mediation of experimenter bias. Journal of Personality, 1963, 31, 313-335.

Shames, M. L., & Adair, J. G. Experimenter bias as a function

of the type and structure of the task. Canadian Psychologist, 1967, 8, 176.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Silver, M. J. Experimenter modeling: a critique. Journal of experimental research in Personality, 1968, 3, 172-178.

Stevenson, H. W., & Hill, K. T. Use of rate as a measure of response in studies of social reinforcement. Psychological Bulletin, 1966, 66, 321-326.

Weiss, L. Effects of subject-experimenter and task variables on compliance with the experimenter's expectation. Journal of Projective Techniques and Personality Assessment, 1969, 33, 247-255.

Wessler, R. L. Experimenter expectancy effects in psychomotor performance. Perceptual and Motor Skills, 1968, 26, 911-917.

Wessler, R. L. Experimenter expectancy effects in three dissimilar tasks. The Journal of Psychology, 1969, 71, 63-67.

Wessler, R. L., & Strauss, M. E. Experimenter expectancy: a failure to replicate. Psychological Reports, 1968, 22, 687-688.

A P P E N D I C E S

APPENDIX A

Low Evaluation Apprehension Communication

Today, you will be helping us to collect some preliminary data which we will use in setting up a subsequent research project. Shortly, you will be assigned to an experimenter who will explain the task to you. In order to make participating more informative and meaningful for you, we will give you a brief description of the purpose of the study.

We are interested in studying people's motor responses. More specifically, we want to find factors (e. g., fatigue, practice, etc.) which affect the individual's responses in various simple manual tasks.

Before we can investigate these different factors, however, we have to know how people perform these tasks when these experimental factors are not present.

That is, we need a control, or standardization, group to use as a baseline against which we can judge the effects that our experimental factors have. This is the reason for your participation today.

We intend to average the performance of all of the students participating today, so that we will have a measure of how subjects perform on the task when such experimental variables as fatigue and prior practice are not present. This information will allow us to judge the effects which our experimental variables have when they are used with a subsequent group of students.

In other words, today's group will help us to find out how subjects typically perform on the task. Later, we can use the data we receive here to judge the performance of subsequent experimental groups of subjects.

APPENDIX B

Medium and High Evaluation Apprehension Communication

Today, you will be participating in a psychological experiment; and shortly you will be assigned to an experimenter who will explain the task to you. In order to make participation more interesting and meaningful to you we want to give you a brief description of the purpose of the experiment. Also, a growing number of psychological researchers are beginning to realize that they have an ethical responsibility to make the purpose of their experiments known to the individuals who are helping them out by participating in their research. As well, we believe that the most valid results will be obtained if you know what is actually happening in the experiment.

We are interested in studying various abilities of people. More specifically, we want to find the factors which improve or impair the various performance abilities of individuals.

Prior research by ourselves and others indicates that, typically, poor performance on these tasks is associated with low intelligence. That is, people who are not able to perform well usually are found to be less intelligent than normal for college students. Much of our initial research in this area indicates that on the basis of performance on simple tasks, we can pick out from a college population those students who would be judged to be of less intellectual abilities.

Several other researchers have presented data which supports the preceding findings. Morgan and Provino (J. of Abnormal and Social Psychology, 1963) for example, report that in a college setting, the Marble Test, which you are about to take, could make rather subtle discriminations between varying degrees of intellectual ability.

The purpose of today's experiment, therefore, is to replicate the previous results, and thus to test further the generality of the finding that people who cannot perform well tend to be of lower intelligence.

APPENDIX C

Post-experimental Questionnaire

1. How concerned were you with your own performance on the Marble Test? Check one.
- +3 Substantially concerned _____
- +2 Somewhat concerned _____
- +1 Neither concerned nor not concerned _____
- 0 Not concerned _____
2. How did the written instructions explaining the purpose of the experiment affect you as you were about to take the Marble Test? Check one.
- 2 Aroused my anxiety to a great extent _____
- 1 Aroused my anxiety to some extent _____
- 0 Had no effect on me _____
- +1 Somewhat reassured me _____
- +2 Greatly reassured me _____
3. Did the experimenter do anything to help you perform better on the Marble Test? ____ If yes, what? _____
- _____
- _____
4. Did you look to the experimenter for clues as to how you might best perform on the Marble Test? ____ If yes, elaborate _____
- _____
- _____
5. What do you think was the purpose of this experiment?
- _____
- _____

APPENDIX D

Mean Response Rates for Subjects tested by
Low-Inducement Experimenters

Expectancy	Trials	Evaluation Apprehension			
		High E.A.	Med. E.A.	Low E.A.	Pooled E.A.
High	1	36.33	34.00	33.75	34.70
	2	36.25	32.58	33.25	34.03
	3	37.17	32.42	33.92	34.50
	4	37.42	32.42	34.08	34.64
	5	38.50	34.08	34.58	35.72
	6	39.00	34.50	35.83	36.44
Low	1	32.67	33.92	34.92	33.83
	2	32.42	33.50	33.75	33.22
	3	33.58	34.17	34.50	34.08
	4	33.83	34.08	33.42	33.77
	5	34.50	34.83	35.83	35.06
	6	35.83	36.08	37.08	36.33

APPENDIX E

Mean Response Rates for Subjects tested by
High-Inducement Experimenters

<u>Expectancy</u>	<u>Trials</u>	<u>Evaluation Apprehension</u>			
		High E.A.	Med. E.A.	Low E.A.	Pooled E.A.
High	1	34.75	33.00	33.33	33.69
	2	35.92	32.17	32.92	33.67
	3	36.83	32.17	32.42	33.81
	4	36.75	32.25	33.33	34.11
	5	38.25	32.75	33.50	34.83
	6	37.92	34.83	34.25	35.67
Low	1	35.92	33.83	34.83	34.86
	2	35.25	32.00	35.25	34.17
	3	34.92	33.08	36.50	34.83
	4	34.67	34.08	36.92	35.22
	5	36.00	34.67	38.08	36.25
	6	36.92	34.58	37.75	36.42

APPENDIX F

Mean Response Rates for Subjects tested by
Intentional-Inducement Experimenters

<u>Expectancy</u>	<u>Trials</u>	<u>Temporal Order</u>		
		Early	Late	Pooled
High	1	34.28	32.17	33.22
	2	33.89	32.33	33.11
	3	34.00	32.44	33.22
	4	34.94	33.00	33.97
	5	35.94	34.89	35.42
	6	36.33	34.67	35.50
Low	1	30.17	33.22	31.69
	2	30.89	34.94	32.92
	3	30.83	34.83	32.83
	4	31.33	35.28	33.31
	5	32.33	36.11	34.22
	6	33.39	37.72	35.56