An investigation of the effect of social desirability on the I-e scale's predictability using the bogus pipeline paradigm.

William Gerald Harris

University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/theses
AN INVESTIGATION OF THE EFFECT OF SOCIAL DESIRABILITY
ON THE I-E SCALE'S PREDICTABILITY
USING THE BOGUS PIPELINE PARADIGM

A Thesis Presented

By

William Gerald Harris

Submitted to the Graduate School of the University of Massachusetts in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August, 1975

Department of Psychology
AN INVESTIGATION OF THE EFFECT OF SOCIAL DESIRABILITY
ON THE I-E SCALE'S PREDICTABILITY
USING THE BOGUS PIPELINE PARADIGM

A Thesis

By

William Gerald Harris

Approved as to style and content by:

Dr. Castellano B. Turner, Chairman of Committee

Dr. Bonnie R. Strickland, Member

Dr. J. William Dorris, Member

Dr. Jerome L. Myers, Chairman
Department of Psychology

August, 1975
ACKNOWLEDGEMENTS

I wish to thank the members of the committee, Dr. Castellano Turner, Dr. Bonnie Strickland, and Dr. William Dorris, for their advice during the preparation of this thesis. I especially wish to express my appreciation to Dr. Castellano Turner, the chairman of the committee, for his support, encouragement, and assistance in the study. I am also indebted to Dr. Bonnie Strickland, for her careful and valuable criticisms and suggestions.

Finally, I wish to thank my wife, Rochelle, for her love, patience, and understanding throughout the study.
# TABLE OF CONTENTS

Chapter I. Introduction .................................................. 1
    Purpose ............................................................................. 11
    Bogus Pipeline Paradigm .................................................. 12
    Specific Aims ..................................................................... 14

Chapter II ........................................................................... 17
    Method ............................................................................. 17

Chapter III. Results ............................................................. 28

Chapter IV. Discussion .......................................................... 39
    Implications ....................................................................... 48
    Summary ........................................................................... 54

References ........................................................................... 56

Appendices ........................................................................... 61
CHAPTER I
INTRODUCTION

For years behavioral scientists have studied how an individual attempts to govern events in his environment. Most of the earlier work on this topic has concentrated on the importance of motivational variables. Alfred Adler (Ansbacher and Ansbacher, 1973) depicted man as being locked into a continual struggle to overcome helplessness. This "striving for superiority" basically derived from what is called man's inevitable feeling of inferiority. Similarly, Frantz Fanon (1963), through clinical observations, inferred that the lack of control over one's environment led to a state of psychological inadequacy. Robert W. White (1959) and Strodbeck (1958) also investigated the efforts of man to regulate his life space. The constructs developed by White described human behavior in terms of "competence and effectance." Several animal studies (Richter, 1959 and Mowrer and Viek, 1948) have attested to the need to be able to effect environmental changes.

However, the effectiveness with which people deal with important events moved away from a strict motivational explanation with the introduction of Rotter's (1966) locus of control construct and scale. Rotter interpreted the control construct as a "generalized expectancy variable" which operated across an array of situations and activities. The Internal-External Locus of Control (I-E) scale, based on the control construct, is a forced-choice measure consisting of 23 relevant items and 6 filler items. The scale's internal locus of control alternative describes people who generally perceive the acquisition of reinforcements
as being contingent upon their own behavior. On the other hand, people with external locus of control scores usually regard the probability of obtaining either positive or negative reinforcements as depending upon outside forces such as powerful others, chance, fate or luck.

Review articles show that research employing Rotter's social learning theory of generalized expectancy and locus of control is proliferate and extensive (Joe, 1971; Lefcourt, 1966; 1972; Rotter, 1966 and Throops and MacDonald, 1971). In many cases the results have supported the usefulness of the scale as a predictor of behavior. Literature on this topic has usually sketched internals as possessing those attributes which led to a healthier, fuller and normal existence while externals are shown to have significantly more signs of pathology.

A brief review of the I-E research will illustrate how it has been employed to assess behavioral differences. One of the many areas being explored ardently by locus of control researchers in the cause of resistance to influence. Crowne and Liverant in 1963 compared subjects in Asch-type conformity situations and founded that subjects high in externality conformed to a greater extent than subjects low in externality. Subsequently, conformers were described as less confident in their endeavors to manipulate important goals in social situations. Likewise, Odell (1959, cf. Lefcourt, 1966) reported findings which showed that externals were more likely to conform than internals. Gore (1963), using an experimenter influence paradigm, also looked at the relationship between locus of control and the resistance to influence. In her study she administered TATs to three groups of subjects to determine which
cards produced longer stories. In the first condition the examiner tried to overtly influence the subjects by revealing which card she thought was the best. The second treatment condition the examiner, through smiles and intonation, attempted to manipulate the subjects. The third condition was a control condition in which the examiner made no suggestions. Gore found no significant differences between internals and externals under the overt suggestion condition and the no suggestion condition. However, when subtle influence was used by the examiner to produce longer stories internals and externals differed in their responses. In particular, internals produced significantly shorter TAT stories than externals. This would imply that internals are less susceptible to subtle persuasion.

Complementary evidence was reported by Getter (1966) and Strickland (1970). Both researchers studied the reactions of internals and externals to social stimuli or influence in verbal conditioning paradigms. Getter's results showed no significant differences occurred between externals and internals during acquisition trials. However, the comparison of subjects during extinction revealed that internals had statistically more conditioned responses in this period. Strickland (1970) noted that denial of having been influenced by verbal reinforcements was related to the locus of control. Specifically, internal subjects denied being influenced by the experimenter and were more likely to follow their own inclinations in regard to giving the correct responses. Whereas, externals when they were aware of the verbal conditioning task exhibited less resistance to the influence of the experimenter.
In another study Phares, Wilson, and Klyver (1971) evaluated the effects of environmental influence in blaming behavior. Subjects in Phares' et al. study were randomly assigned to either the non-distractive condition or the distractive condition. Afterwards they were given tests that supposedly measured intelligence. Instead the tests were designed to yield results characteristic of a poor performance. The findings define internals as being more reluctant to use blaming behavior than externals across both situations. In fact, internals blame attribution showed a direct relationship with increases in confusion and noise.

Yet research concerning cigarette smoking has shown that internals are not simply obstinate and unyielding to external influence. James, Woodruff, and Werner (1965) reported that following the U.S.P.H.S. Surgeon General's report linking cancer with cigarette smoking, among male smokers, those who quit and did not return to smoking in a specified period of time were more internal than those who believed the report, but did not stop smoking. Using role-playing procedures, Platt (1969) carried out an investigation in which subjects role-played as physicians, patients, or observers during a medical examination report containing bad news for the patient regarding cancer and smoking. Platt reported more success at influencing the smoking behavior of internals than of externals.

Differences in how internals and externals resist influence seemed to suggest differences in their cognitive activities. Seemingly, persons with internal control expectancies are more curious and more likely to
take advantage of situations that persons with external control expectancies. Lefcourt also viewed cognitive activity as a function of locus of control and stated that:

to maintain a generalized expectancy of internal control should require some modicum of success at steering oneself around obstacles and toward desired ends. A number of invalidations, or negative reinforcements, should serve to increase self-doubts and consequently lead to increasing external control expectancies. Such self-direction should entail more active cognitive processing of information relevant to the attainment of valued ends, and should be reflected in the types of cognitive strategies that come to characterize the person (Lefcourt, 1972).

Two of the earliest studies of the relationship between cognitive activity and locus of control were undertaken by Seeman and Evans (1962) and Seeman (1963). Both investigations reported differential learning between internals and externals in two field settings. Seeman and Evans used 43 pairs of white male patients, with each pair matched on demographic and hospital variables. They concluded that hospitalized tuberculosis patients with internal control expectancies had obtained more information relevant to their personal conditions. Externals, on the other hand, were less knowledgeable about their condition and questioned the hospital staff less about health matters. In the second study, Seeman (1963) demonstrated that reformatory inmates, who were internal had secured significantly more information about successfully achieving parole than inmates with external control expectancies. Also, when information was less salient from a personal standpoint internals and externals did not differ in their retention of the material.
Several other studies have endorsed Seeman and Evans (1962) and Seeman (1963) major conclusions that internally control persons often times perceived and value information differently than externally control persons. Davis and Phares (1967) noted that internals actively sought more information relevant to the task of influencing the attitude of a person towards the Viet Nam war. In another publication, Phares (1968) compared how internals and externals utilized information in decision-making processes. Phares discovered that after one week internals in this computer-simulated task were more likely than externals to make use of information which was equally available to both. Lefcourt and Wine (1969) observed the way in which internals and externals attend to social cues while trying to learn about another person. Each subject in the experiment conducted interviews with two distinctly different target persons. One of the target persons presented himself to the subject in a perplexing manner, avoiding eye contact while the other person was more conventional with regard to eye contact. Lefcourt and Wine found that internals tend to look at the eye-contact-avoidant person much more frequently than they looked at the conventional target person, and they displayed more curiosity and attentiveness toward the eye-contact-avoidant person than did subjects with external control expectancies. Furthermore, internals as evident by their looking behavior also made more observations of both target persons. From these findings Lefcourt and Wine inferred that internals are more effective than externals in focusing on informative cues in their social environment.
Research with the locus of control construct is broad and voluminous. Some of the other pertinent areas of investigation include deferred gratification, achievement behavior, and the response to success and failure. Lefcourt reported in his 1971 research report that locus of control and reinforcement preferences are related. In short, he suggested that delayed reinforcement schedules were preferred by internals because they (internals) are more accustomed to long-term efforts directed toward distant, but valued ends while on the other hand impulsivity and immediate rewards appear characteristic of externals. Achievement behavior has also been linked positively to a sense of personal control (McGhee and Crandall, 1968). Externals, however, showed less vitality in their academic pursuits. This is understandable since externals generally do not perceive rewards or punishment as being consequences of their behavior. Likewise, research has explored and demonstrated differences in the way in which internals and externals respond to success and failure experiences. Lefcourt, Lewis, and Silverman (1968) found that internals performed better than externals under conditions where skill controlled the outcome, while externals performed better than internals in chance-determined conditions. Rotter and Mulry (cf. Rotter, 1966) studied how internals and externals respond to an angle-matching situation of extreme difficulty. They observed that internals took longer to decide on a matching standard under skill conditions than did externals, but took a shorter time under chance conditions than did externals. As Rotter and Mulry suggested internals appear to value reinforcements that were contingent on skill more than chance and vice versa for externals.
Several extensive review articles have reported that the test-retest reliabilities of the I-E scale are acceptable (Lefcourt, 1966; 1972; Rotter, 1966). Reliability scores for various sample populations and time intervals have regularly fallen between .49 and .89. Hersch and Scheibe (1967) also looked at the reliability and validity of the I-E control scale as measuring a personality dimension. Hersch, et al. outcomes furnished additional evidence supporting the relative merits of the construct. Specifically, they found that the I-E scale consistently measured maladjustment with internals less maladjusted. Also, they established that the I-E scale is related to some personality scales. Duke and Nowicki (1973) in a partial replication of Hersch's and Scheibe's study administered a new measure of I-E, the Nowicki-Strickland I-E Scale for Adults, and obtained analogous results. And finally, Harrow and Ferrante (1969) working with 86 patients generated a test-retest reliability over a six week period of .75.

Nevertheless, Rotter's generalized expectancy of reinforcement principle is not completely free of controversy and debate. In fact, several studies have seriously questioned the predictive power of the I-E scale. Levenson (1973), in an investigation of the I-E dimensionality obtained results that conflicted with the unidimensional interpretation set forth by Rotter (1966). Using Kaiser's Varimax method and a sample population of functional psychotic and neurotic inpatients, Levenson revealed that three factors instead of one were embedded in the I-E structure. These three distinct aspects of locus of control are called: internality, control by powerful others, and chance dimension
were moderately high while the internality scale's reliability was extremely low (Levenson, 1973). In general, it appeared that the powerful others and chance scales produced meaningful and consistent descriptions for maladjusted persons. On the other hand, the internality factor seemingly reflected a momentary perceived locus of control. Similarly, Collins (1973) analyzed the Rotter I-E scale and found four separate relatively orthogonal subscales. Specific components identified by Collins were: belief in a difficult world, a just world, a predictable world, and a politically responsive world. In addition, Abramowitz (1973), Gurin, Gurin, Lao, and Beattie (1969), Joe and Jahn (1973), Mirels (1970), and Lao (1970) strongly supported the notion that the locus of control construct was multidimensional. At the present time, however, it is difficult to assess completely the value of the unidimensional-muti-
dimensional argument since most psychological tests and scales have at one time or another been criticized about their capacity to consistently measure a specific variable.

Still, a seemingly inherent problem area is the I-E scale's ability to distinguish personality factors from social norms (e.g. social desirability). Nowicki and Hopper (1974), and Palmer (1971) reported that alcoholics, who were characterized by their dysfunctional behavior and impaired emotional state, manifested significantly more traits in the external direction. However, Goss and Morosko (1970) presented disconfirming data. In Goss's et al. study alcoholics scored in the internal control direction. Furthermore, internally-oriented alcoholics indicated less anxiety, depression, and clinical pathology on the MMPI.
Distefano, Pryer, and Garrison (1972) tested alcoholics and emotionally impaired patients with Rotter's I-E scale. Distefano's et al. results showed alcoholics to have strong beliefs in personal control. Finally, Berzins and Ross (1973) compared 600 hospitalized opium addicts to 800 college subjects. Their conclusion endorsed Goss et al. and Distefano et al. findings of internality for addictive populations. The drug addict's incessant abuse of drugs, they surmised, induced a generalized belief that he (the addict) can control salient reinforcement. This "pseudo-internality of addicts" was apparently a consequence of the drug effects which liberated the addict from uncontrollable feelings (e.g. anxiety, moods, impulses and other forms of distress).

Most of the I-E literature has explained the internal control as comprising a more homogeneous grouping than external alternative. Still some researchers have advocated the redefining of the construct's categories (Hersch and Scheibe, 1967; Ross and Berzins, 1973). Additional clarification of the I-E's alternatives, they insist, would multiply the scale's worth and effectiveness in clinical settings. Yet few studies have explicitly considered the impact of social desirability on the individual's response set to the I-E scale. Even though few studies have examined the influence of socially acceptable responding to the I-E scale, Rotter and his colleagues were sensitive to this problem during the scale's construction. For example, they selected a forced-choice questionnaire format as one method of controlling social desirability. In another effort to minimize social influence, they reduced and purified the scale from its original 100 forced-choice items.
to 60 items and finally to its present version of 23 relevant items. This was done by eliminating those items which either had a high correlation with the Marlowe-Crowne Social Desirability (SD) scale (Crowne and Marlowe, 1964) or a correlation approaching zero with its validation criteria (Rotter, Liverant, and Crowne, 1961; Seeman and Evans, 1962).

Unfortunately, these safeguards alone are not able to manage the effects of socially acceptable responding since how a person reports his perceived locus of control might possibly be confounded with his fantasy of what is good, right or expected. As briefly stated above, Rotter in his 1966 review article acknowledged the fact that the I-E scale was not entirely free of social desirability. Moreover, he cited several studies in which the I-E scale and the SD scale were correlated. These various correlations showed an average relationship slightly higher than -.20. Likewise, Altrocchi, Palmer, Hellman and Davis (1968) and Feather (1967a) respectively computed significant relationships between the I-E and SD scales of -.34 and -.42. These moderately high correlations attested to the difficulty in trying to lessen social attractiveness in response patterns. Indeed, such statistically relevant correlations question the discriminatory validity of the I-E scale, which brings us to the purpose of the present study.

PURPOSE

It was the intention of this study to examine whether variable confounding played a significant role in the inconsistency of I-E find-
ings. As alluded to earlier, attempts to measure the subject's real attitude in an experimental setting may often become adulterated with transient, socially attractive types of responses. This suggested that manufactured answers should be suspected when systematic replication studies give conflicting results. I-E studies using alcoholic populations provided an illustrative example of a potential lack of authenticity in response patterns. Findings have favored both an external locus of control (Palmer, 1971) as well as internal expectancy of reinforcement (Distefano et al. and Goss et al.) interpretation.

Also, the study looked at the utility of the Bogus Pipeline Paradigm as a real attitude indicator. However, before turning to the specific predictions, a brief review of the bogus pipeline research will be considered.

**BOGUS PIPELINE PARADIGM**

The bogus pipeline, a new rating procedure, has been offered as an improvement on previous rating measures (Jones and Sigall, 1971). This methodological device supposedly measures one's real feelings about a person or an issue. The acquisition of uncontaminated data from subjects was accomplished by persuading them (the subjects) that the electro-physiological equipment calculated certain uncontrollable bodily reactions. Thus upon being exposed to a convincing demonstration of the machine's ability to measure attitudinal direction and intensity precisely, subjects are then requested to predict the machine's readings (a dependent variable). This strategy incorporated a generally well-accepted myth that "truth-machines" are reliable sources of the truth. In
fact, Ostroms (1973) questioned this lie detector-like approach on ethical and practical grounds, but such criticisms have not received additional support. Obviously, in the field of psychology deception techniques are standard procedures in many experimental situations.

Valins (1966) conducted a false feedback study in which subjects viewed 10 slides of seminude females while listening to bogus heart beats. Half of the subjects heard their heart rates increase markedly on 5 of the slides and no noticeable variation to the other 5. Valin's findings indicated that those stimuli associated with heart rate changes were rated significantly more attractive.

In another instance, Sigall and Page (1971) examined response distortion in a study of stereotypes. Subjects were instructed to rate 1 of 2 ethnic groups. Negroes or Americans on how each of 22 adjective traits characterized the group. One group of subjects responded by predicting their electromyograph readings whereas the second group received standard rating forms. In comparison the two groups showed meaningful differences, with subjects in the electromyograph (EMG) condition exhibiting more negative responses. For example, the trait, honesty, was depicted as more characteristic of Negroes than Americans in the rating condition, but less characteristic in the bogus pipeline condition. Also, Americans registered more favorable responses in the EMG (the bogus pipeline situation) than in the rating condition. These conclusions lend support to the idea that the bogus pipeline procedure reduces social desirable responding.

The two aforementioned bogus pipeline experiments illustrated
different versions. Sigall and Page (1971) utilized the paradigm as a dependent variable measure while Valins (1966) employed the bogus pipeline system as an independent variable. This project has adopted the bogus pipeline as an independent measure.

**SPECIFIC AIMS**

The aim of the study was to look at the extent to which socially desirable responding may obfuscate the scale's predictability. This task demanded an experimental design capable of inspecting a variety of important and related hypotheses. Therefore, the following hypotheses attempted to examine and evaluate the effects of treatment and social desirability on I-E outcomes. Furthermore, the treatment manipulations should cause specific changes in the social desirability factor. In particular, this factor should be reduced, normalized or increased. It was presumed that the changes would correspond respective-ly to these treatment conditions: the bogus pipeline, the conventional, and the validation.

1. As mentioned briefly above, it was predicted that the amount of social desirability embedded in response sets to the I-E scale would vary with respect to the treatment conditions. Accordingly, the bogus pipeline treatment condition was structured so as to reduce the possibility of subjects responding in a socially acceptable way. Studies by Sigall and Jones (1971) and Sigall and Page (1971) have reported that a subject's response pattern was less influenced by socially attractive forces when he believed that his responses were being monitored and
analyzed by electronic equipment. The conventional condition, also called the control condition for this experiment, was merely the standard method of using the I-E scale. Under this setting it was proposed that social desirability responding would be greater than in the bogus pipeline treatment. This hypothesis was based on the presumption that subjects not inhibited by explicit manipulations from responding in a socially attractive way are more likely to display socially desirable responding patterns. The validation condition of this experiment offered an adequate method of documenting the amount of social desirability responding in the other conditions. In this situation subjects were instructed to respond in the most socially acceptable way to the I-E scale. This hypothesis in essence expected social desirability responding under the validation condition to be increase significantly. Thus it was hypothesized that social desirability responding would be markedly greater for the validation subjects than for subjects assigned to either the bogus pipeline or the conventional setting.

2. Central to the aim of the study was the prediction that bogus pipeline internals would be less internal than their counterparts under the conventional condition. This was based on the fact that social desirability tendency in the bogus pipeline treatment has been reduced; whereas in the conventional situation this social force was operating at its normal level. Rotter
(1966) observed that correlations between the I-E scale and the SD scale tended to be on the average slightly higher than -.20.

In addition, this study investigated whether there existed on the social desirability parameter any real differences in how males and females responded to the I-E scale. Eisenman and Platt (1968), Gore and Rotter (1963), Hamsher, Geller, and Rotter (1968), and Rotter (1966) concluded that there was no appreciable sex differences on the I-E scale. However, these results have been attacked most notably by Feather (1968) who reported that females earned significantly higher external scores than males. Therefore, the relationships amongst sex, I-E scores, experimental situations and social desirability was explored.
CHAPTER II

METHOD

SUBJECTS AND OVERALL PROCEDURE

Subjects for this study were recruited from predominately large lecture classes in psychology. In soliciting subjects the experimenter gave an in-class presentation, describing the purpose of the experiment. The explanation suggested that the study was designed to examine certain salient parameters and characteristics of attitude formation. Students were further notified that the study entailed a follow-up phase and that all students deciding to participate in the initial part, should also be willing to complete the second part of the task.

The incentive for student participation was in the form of experimental credits. A total of three experimental credits could be earned by participating in the study. Such credit could be used towards the student's final grade in the course. If for some reason, the student was unable to participate in the second phase, only one experimental credit was offered. Furthermore, to help insure that subjects were involved with both aspects of the project, credit was awarded, when possible, upon completion. This approach was undertaken in hopes of reducing absenteeism and withdrawals from the experiment. Although the withholding of experimental credit aided in limiting the number of no-shows, it was also used to provide some control over the time interval between testing since individuals involved in the follow-up who failed to show for the second time or were unable to find a convenient time within an allotted time period were replaced and given only one
experimental credit.

In addition, the use of a two stage task meant that subjects were not allowed to remain anonymous. Personal information, such as name and telephone number, had to be obtained from each respondent for the exclusive purpose of contacting subjects involved in the latter half of the experiment. However, each subject was informed at the beginning of the study that all information would be transferred to a coding system, as soon as their role in the project had been finalized. More importantly, assurance was given to them that the background data requested and their responses to the items on the scales would remain completely confidential. The presumed effect of conveying such information to subjects was to dispell, in part, apprehensions which might restrict honest and free reactions from participants.

After hearing the intention of the study, students who volunteered were grouped together in the lecture room setting and instructed to fill-out a social reaction and a personal reaction inventory, commonly known as the Internal-External Locus of Control scale and the Marlowe-Crowne Social Desirability scale respectively (see appendices A thru C). Briefly, the underlying constructs of the I-E and SD scales served to delineate subjects on two basic personality variables--internal vs. external locus of control and high vs. low need for social approval. Yet the subject's status on the scales is relative to the distribution of scores of the population sampled. Consequently, no fixed or static cutoff level can be readily used across different subsets of people drawn from the general population. Instead a median split was done independently on
each set of scale scores. This in effect furnished the necessary grouping criteria. For instance, the median split for the distribution of internal-external scores occurred at 11. Subjects with scores below 11 were classified as having an internal locus of control of reinforcement. While on the other hand, participants whose generalized expectancy score fell in the range from 11 to 23 were labeled as externals. On the SD scale, the bisection of the distribution of scores was at 13. A person with a low need for social approval had a score within the zero to 13 range; a high need for social approval encompassed subjects whose scores fell in the 13 to 33 range.

Subsequently, each subject's group assignment depended on the subject's outcome on the two inventories. In other words, dividing the list of scores on each scale at the 50th percentile created four subcategories: internal high, internal low, external high, and external low. In this dichotomizing portion of the study, it became necessary to examine a fairly large sample population before each subgroup was amply represented, both in terms of size and sex. The external high subgroup was by far the most difficult to fill. This may be indicative of a characteristic of college students, whose achievement orientations make it more likely for them to believe that they have some control in mapping their destinies. The search for students with external high scores required the sampling of over 500 people before sufficient number was obtained.

Out of this large potential experimental population 255 individuals completed the experimental phase. However, only 247 subjects were analyzed. The remaining 8 subjects were not included because pertinent
data were either missing on them or their age deviated significantly from the population mean.

In the dichotomizing phase of the study subjects were also asked a set of background questions. This selective inquiry helped to elucidate the characteristics of the sample. For instance, a 123 female and 124 male subjects comprised the sample population. Ranking on the basis of years in school saw the sample spreading itself fairly evenly over the various classes. The population's mean GPA of 3.1 was consistent with the University's undergraduate community. Subjects reported areas of resident as follow: 31% small town, 32% town, 32% city, and 5% rural and unspecified areas. The academic majors of the individuals provided still another salient view of the sample. 37% reported psychology as their major while 47% of the subjects stated that their academic discipline was not psychology. Subjects still undecided on their primary area of interest totaled 15%. Finally, the mean age for the population was 20.1, with a range of 17 to 26 (see Appendix F for a copy of the Personal Data Form).

**EXPERIMENTAL DESIGN**

To test the specific hypotheses of this study, a randomized multi-factor 2 (internal/external) X 2 (high-low/SD) X 2 (sex) X 3 (treatment) design with an alpha-level of .05 was chosen. This design was the most efficient because within it each of the specific hypotheses was readily testable.

In addition, an eight subject per cell minimum was strictly enforced.
This was done because it was necessary to maintain sufficient power throughout the design if the conclusions reached about treatment effects were to be justifiable.

**PERSONALITY MEASURES**

The Marlowe-Crowne Social Desirability scale was used to assess the need of the subject to respond in culturally sanctioned ways. This 33 item, true-false inventory contains statements about behaviors of high social approval and appeal. Accordingly, the degree to which an individual responds to these descriptions in a socially desirable fashion provides a measure of the tendency to depict himself "in improbably favorable terms" (Marlowe-Crowne, 1965, p. 39). An illustrative item is: I would never think of letting someone else be punished for my wrongdoings (see Appendix C for a complete list of items).

The Internal-External Locus of Control scale is a 23 item, forced-choice scale measuring the extent to which a person attributes the locus of control of events in his life to himself or beyond his personal control. Each item pairs an internal with an external choice in the following manner: I more strongly believe that (a) I have usually found that what is going to happen will happen, regardless of my actions (b) trusting to fate has never turned out as well for me as making a decision to take a definite course of action (see Appendix A for a list of the items).

**PROCEDURES**

As mentioned earlier, students at the beginning of the semester were administered the two personality inventories. These two indicators
served as independent variables and also functioned to stratify the population. After designating the subject's group type, each respondent was then randomly assigned to an experimental condition.

Three experimental conditions—the bogus pipeline, the conventional, and the validation—were employed in the second stage of the project. In this second stage the dependent measure was the I-E scale. The following is a description of the experimental manipulation procedures used in the study.

CONVENTIONAL CONDITION

The conventional method and the initial screening and identification phase of the study were similar. In both cases the instructions used with the Social Reaction Inventory (the I-E scale) were those specified by Rotter (1966). However, the conventional method differed from the initial screening phase because it only employed this one inventory. In essence, this method functioned principally as the control condition for the experiment since it was merely a re-administration of the I-E scale.

VALIDATION CONDITION

The validation and conventional methods differed only in instructions. In the conventional procedure, the I-E scale measures one's personal belief; whereas, the I-E scale under the validation method assesses social belief. That is, the subject in this condition selected the one statement of each pair which he/she strongly felt to be the most socially desirable alternative. This in turn provided an effective way of documenting social desirable responding in the other experimental
conditions. The other advantage was that it allowed us to look at the
direction of social desirable responding and whether this direction was
consistent for internals and externals (see Appendix B for instructions
to the I-E scale).

THE EMG CONDITION: THE BOGUS PIPELINE TECHNIQUE

The introduction of the EMG was the most intricate. Subjects,
randomly assigned to this experimental situation reported individually
to an experiment entitled "Attitude Formation Study." Upon arriving,
each subject was presented with a preliminary 4 item, yes-no question-
aire which was completed prior to entering the experimental room
(see Appendix D). Furthermore, these items dealt with issues of a non-
sensitive nature since it was important to minimize the possibility of
response shifts. The purpose of the questionnaire was to obtain in-
formation that would be used to demonstrate the accuracy of the electro-
myograph machinery to predict a subject's true attitude or belief. After
completing the questionnaire the subject was told to place it in a box
on top of the desk. The subject was then invited into the experimental
room by the experimenter who was an advanced graduate student in clinical
psychology. At this time, a confederate housed in an adjacent room emerged
and reproduced the subject's responses.

In the experimental room the subject was seated in front of an
electronic console. The equipment was described and fully defined to him
by the experimenter as an adapted EMG that is able to measure "implicit
muscle potential" (consult Appendix D). Based on this premise, the
experimenter insisted that no verbal activity of the subject was needed to predict his real attitude on any topic. However, the experimenter did concede that the apparatus occasionally needed to be adjusted to operate accurately. He then proceeded with the help of the subject to readjust the EMG’s baseline. The experimenter instructed the subject to listen carefully to each question, but not to verbalize his answer. Coincidentally, the experimenter pointed out a portable meter box located on top of the console to the subject which registered the subject’s reaction after the question has been read.

Actually, the confederate in an adjacent room who earlier copied subject’s responses to the items on the preliminary questionnaire, was regulating the direction of the meter pen. The subject was unaware of this deception. The common reaction amongst subjects was one of amazement and/or apprehension. Moreover, most people on the post-questionnaire indicated that they were convinced of the EMG’s power to detect their real attitudes.

Upon demonstrating the EMG’s effectiveness, the experimenter stated that he wanted the subject to try to predict which one of the two choices the EMG would register as more characteristic of him. A cassette recording of the I-E scale items was presented and the subject was asked to choose the one most likely chosen by the EMG. Afterwards, the subjects were given the post-questionnaire to fill-out and encouraged to express their opinions about the experiment. Finally, each subject was informed that a debriefing statement would be sent to them by mail (for the complete introductory statement to the EMG as well as the letter which
was sent to all subjects, see Appendices D and E).

**POST-QUESTIONNAIRE: CONFIRMATION OF THE BOGUS PIPELINE TECHNIQUE**

In this study the authenticity of the bogus pipeline paradigm was examined with the aide of a post-questionnaire. The basic objective of the questionnaire was to ascertain how subjects perceived this experimental situation. Obviously, without undertaking some systematic assessment of subjects' perceptions of the study's purpose and their impressions of the electronic apparatus, conclusions would be tenuous. Therefore, a 4 item, open-ended questionnaire was composed to check subjects' perceptions of the bogus pipeline setting. The foundation of the questionnaire was based on several salient questions. For example, did the apparatus have an authentic appearance? Was the experimenter's description of the machinery's function believable? And was the pre-test for baseline adjustment convincing enough to the testee to instill a belief in him that the equipment could measure his real attitude in a variety of situations? Although these are important questions, they were felt to be too leading and consequently more incline to heighten suspicion in subjects. The terminology was instead changed, without disturbing the essence of the questions. Subjects were thus asked to describe the purpose of the experiment and to indicate whether or not they felt this method was an effective way of securing a person's real attitude. Requesting subjects to briefly explain the experimental objectives was done to see if they have fully digested the given description. Likewise, confirmation of the procedure for eliciting a real response from an
individual was taken to mean that the subject was influenced by the pre-test demonstration and believed that the equipment had a genuine use.

An overall look at the bogus pipeline sample group showed that 69% of the population endorsed the technique. Most of the subjects' endorsements were clear acceptance of the apparatus's power. One subject, for instance, said that this procedure was a useful mode of obtaining a person's real attitude because "an individual is more likely to be honest (if) he thinks his physiological reactions are being recorded." Another respondent remarked "coupled with one's fear of being outsmarted by the EMG machine one will try to answer spontaneously and honestly."

A more amusing but still confirming statement comes from yet another subject: "It strikes me that you would not be able to hide much from 'George,' the machine." Finally, a student noted "(that) given the comparison at the beginning of the experiment with the readings on the EMG and my responses on the questionnaire (i.e. pre-test), it works."

These are typical responses of subjects who approved of this method of assessing real attitudes. (Consult Table 1 for a breakdown of subjects' views of the EMG procedure.)

Subjects not in favor of employing the bogus pipeline procedure to get at one's real attitude seldom questioned or suspected the accuracy of the equipment; instead they frequently focussed their arguments on the scale. It was common to find subjects, who rejected the technique, complaining about the poor match between alternatives of the items. As one subject put it, "the questions used were very ambiguous and it was hard to say whether one response was really better than the other."
Another subject stated that "neither choice applies to the way I feel, even though my physical reactions may show it." The general sentiment of the 20% who disapproved seemed to be as one subject said "that the choices are too one-sided." In fact, only two subjects or 3% of the bogus pipeline's total sample made comments skeptical of the experiment's intention or the pipeline's purported ability. The lack of overt suspicion among subjects in the bogus pipeline situation strengthens the conclusions inferred from the data.

When the bogus pipeline sample was divided into its internal and external categories, several striking differences were noticeable. Internals, on the whole, were far more likely than externals to support the electronically controlled method of measuring someone's real attitude. 78% of the internals considered the bogus pipeline system effective, whereas, only 58% of the externals subscribed to this procedure. Furthermore, only 14% of the internals in the present study criticized the technique used to extract a true response from a subject while the polling of externals showed that 2% opposed it.

Similar findings appeared when the subject's sex was also taken into account. Internal females tended more than their male counterparts to approve the implementation of electronic equipment as a way of gaining access to a person's real feelings. On the other hand, male and female externals differed only slightly in their view of the situation.
CHAPTER III
RESULTS

RELIABILITY OF SCALES

Internal consistency estimates for both I-E scale and the SD scale were obtained by using the Cronbach Alpha. Cronbach Alpha is a generalization of the Kuder-Richardson formulas 20 and 21. In short, "Alpha" assesses the degree to which items are independent measures of the same construct (Bohrnstedt, 1972).

In the present study, the I-E and SD scales in the identification phase of the experiment registered Alpha reliabilities of .72 and .74, respectively. The moderately high reliability estimate of the locus of control scale is comparable to reliability data reported in Rotter's (1966) review article for groups made up of males and females. Although Rotter employed the Kuder-Richardson method of calculating internal consistency instead of Alpha, both techniques are conceptually equivalent, the difference being that the Kuder-Richardson (1937) formulas are restricted to use with dichotomous items; whereas Alpha also has the ability and power to accurately examine polychotomous items.

An internal consistency coefficient of .88 for the SD scale was reported by Marlowe-Crowne (1960) who applied the Kuder-Richardson formula 20 to their data. The .74 coefficient found in the present study, although quite satisfactory, falls below this earlier finding. Several related explanations for this discrepancy in the reliability measurement are suggested here. First, Marlowe-Crowne's reliability coefficient was computed on a small sample, totaling only 39 subjects.
Moreover, a disproportionate number of males and females were represented in the study. In particular, 29 females and 10 males were chosen, which suggest that the overloading of females may have increased the reliability. Additionally, the mean age of the Marlowe-Crowne sample was 24.4 years. Comparison of these features with similar features in the current project reveals striking differences. A considerably larger number, 247 subjects recruited, presumably had an impact on the reliability estimate. The sample furthermore had an equal proportion of male and female respondents and their mean age was just 20.1 years. These characteristics of the sample may have been in part responsible for driving the internal consistency reliability of the SD scale down. Yet, even though the SD reliability departs from the stronger result of the previous research, the computed coefficient of the SD scale like the coefficient found for the I-E scale in the identification stage is clearly acceptable.

Internal consistency analyses were done separately for each treatment group (see footnotes for Tables 3-10). This was done to provide an even more exacting and clearer picture since both testings of the I-E scale, along with the SD scale were considered. When viewing the reliabilities for the bogus pipeline, for instance, the I-E scale from the first to the second administration showed a slight decrease in the internal consistency of the scale. This finding implied that to some extent the dependability and predictability of the scale's items have been lessened, though only slightly. More importantly, the unique environmental situation used to conduct the second testing of the I-E scale ap-
peared to have disturbed the stability of the instrument. However, the other two experimental groups, the conventional and validation, displayed increases in their internal consistency estimates during the second testing of the locus of control scale. Specifically, the increase of the conventional group was small and appeared to be merely a function of practice since subjects were already familiar with the items. On the other hand, the validation group showed a marked increase in its reliability estimate for this second presentation. This improved coefficient was probably due partially to a practice effect. Moreover, in this condition subjects received instructions which explained how they should respond to the items on the scale. Consequently, homogeneity among subject's responses was heightened and thereby reducing variability because subjects were informed to react to the items in the most socially desirable way. Furthermore, as a homogeneous population, college students may tend to have similar views on what is denoted as carrying society's approval.

To get a further indication of what the scales in this investigation were measuring, an item analysis was carried out. (For a fuller presentation of the correlation matrix consult Tables 3-10.) The function of this statistical procedure, in this instance, was to help the researcher to get a clearer impression of the goodness of fit amongst items of the scale and also how the total score related to the various scale items. Item analysis is thus an excellent means of checking the dimensionality of the items in terms of whether they are measuring the same construct. This was conducted by examining the outcomes of three types of correla-
tions. First, all possible item combinations with another item were correlated. The information acquired from the resulting correlation matrix facilitates the detection of suspicious trends which might raise questions concerning the homogeneity of the items. The other type of correlations were done between the item and the total score for the scale. In particular, the item could be included or excluded from the total score. When the item being paired with the total score has not been removed from the latter, the correlation is defined as an uncorrected correlation. Some social science researchers look upon the inclusion of the item in the total scores as causing confounding and a spuriously high correlation coefficient. This is quite understandable since embedded in the overall coefficient of the item with the uncorrected total score is the correlation of the item with itself, which yields a correlation of 1.00. The corrected correlational findings for the overall population on the first I-E testing are comparable to Rotter's (1966) published results.

In addition, test-retest information on the I-E scale was gathered for each treatment group (see Table 12). The findings showed the computed reliabilities for the bogus pipeline and conventional situations to be inflated. The short time span between testing was the likely reason for these large coefficients. Indeed, the time between the two testings seems to be a salient parameter in any attempt to explicate the reason for the moderate increase since the mean time interval was only fifteen days. Rotter (1966) demonstrated that the test-retest reliabilities may actually decrease significantly when a period of two months is required between
testing. Of more importance to this study was his results that attested to the fact that a smaller time interval tends to induce an increase in the test-retest reliability estimate.

On many occasions the I-E scale and SD scale have been used together and it is now very common to study the relationship between these two scales (Rotter, 1966). Their correlations, as reported by Rotter (1966), tend to range from -.07 to -.35. The median correlation of -.22 represents different college populations of male and female students.

In the present report an equal number of male and female college students participated and a correlation of -.15 was found between the first I-E and SD scale scores. Breaking the sample population down into treatment conditions revealed that the locus of control of the identification phase and the SD scale were inversely related. In fact, these coefficients went from a -.10 to -.20, well within the previously reported range.

Examination at the treatment level of the experimentally employed locus of control scale with the SD scale resulted in zero-order correlations for both the conventional and validation subgroups. While the I-E of the bogus pipeline with its corresponding SD results produced a coefficient of -.26, by far the strongest correlation.

ANALYSIS OF VARIANCE: TEST OF MAJOR HYPOTHESES

A multivariate statistical procedure (Finn, 1972) was used to inspect the major hypotheses of the study. This computational method was deemed the most efficient because it was able to accommodate unequal cell sizes and randomized factors. Moreover, the formulated hypotheses of the
project concentrated on how the dependent variable appeared when examined in terms of an independent variable or first order interaction between independent variables. Subsequently, a series of main effects and interactional hypotheses were analytically scrutinized. However, before focusing on these postulations it should be reiterated that the dependent variable is the outcome on the second testing of the I-E scale, while the independent factors selected for analysis are the treatment groups, the I-E score in the pre-test, the SD score and sex.

Table 13 is a presentation of the results of the analysis of variance, including main and interaction effects. Probably the most critical hypothesis of the study was the main effect due to treatment conditions. Recalling the specific aims, described in an earlier section, it was proposed that social desirability responding would be a function of the treatment conditions. Specifically, the hypothesis contends that social desirability would be less evident in the bogus pipeline procedure, prominent in the conventional setting, but greatest in the validation method. The validation situation in this particular instance was also utilized to determine precisely the direction of the socially acceptable responding in our population sample. The test of the treatment main effect was significant (F=3.85, p .05). This unmistakly conveyed the notion that the various experimental situations made a difference in how subjects reacted to the second I-E testing. An examination of Table 17, the mean table for the treatment groups, showed as expected, that the mean of the bogus pipeline condition was the most external. In accordance, the validation sample recorded the lowest mean score and represented the most internal condition; while the mean score for the conventional situation fell between the bogus
pipeline and the validation groups.

As expected, a highly significant main effect occurred for the first locus of control scale presented to subjects (F=52.77, p .001). This suggested that a strong relation existed between the scores of the two administrations of the I-E scale. A significant main effect due to SD level was not obtained. This was, however, expected since the primary role of the scale was simply to control for the potential effect of social desirability level. The main effect for sex was found to be not significant.

First, second, and third order interactions supplied still another body of information. The first order interaction of the treatment condition with the first set of responses to the I-E scale proved significant (F=22.61, p .001). The interaction of these two independent variables is related to variation in the dependent measure. The influence of the interaction of the treatment groups with the SD scale on the dependent measure was beyond the level of significance defined for the study. Similarly, the relationship of the treatment group with sex did not produce a statistically significant F-ratio. Furthermore, the interactional hypothesis between the overall mean of the first I-E score and the mean score for the SD scale was not confirmed. Also, the interaction of the first I-E testing with sex yielded non-significant results. Finally, the combined effect of the SD level and sex on the dependent measure produced a probability score greater than the .05 level of significance agreed upon for this project.

The effects of three and four variable interactions for the most part came out non-significant. The only significant F-test was one in
which the interaction consisted of the treatment groups, the first I-E testing and sex (F=3.09, p .05).

Inspection of the relationship of treatment groups, the first I-E and the mean of the SD scale was not statistically significant. Nor was the significant test of the second order interactions of treatment groups, SD and sex in accordance with the study's .05 level of significance. Thus, the non-significant results indicate that the interactions did not differ appreciably from chance. A third order interaction was also investigated. This interaction takes into account all four independent variables. Triple as well as higher interactions seldom yield significant F-tests and the study's only triple interaction was likewise non-significant.

**T-Tests**

The principle usage of t-tests was to confirm and to clarify the meaning of any significant results found in the analysis of variance. Obviously, this is a very complementary procedure since pairwise examination of both treatment means and cell means was conducted. The ANOVA described above checked the variability across the means. The t-test, though limited to examining the differences between pairs of means, can also take into consideration the direction of the difference. Thus, because the direction of the difference was important, a one-tailed test of significance was used with alpha set at .05.

As previously mentioned, the resulting outcome of the treatment means in terms of their magnitude and direction was anticipated. The t-tests indicated that only the means of the bogus pipeline and the validation
condition were significantly different. (Consult Table 19 for a complete examination of t-test results for treatment conditions). This is quite reasonable since the bogus pipeline and the validation categories were designed to produce opposite effects. That is, the bogus pipeline subgroup sought to extract the subject's truest attitude towards items on the I-E scale; whereas the validation situation instructed subjects to respond in the most socially desirable way.

Using this parametric statistic to analyze the difference between the bogus pipeline and the conventional treatment yielded a \( t \) of 1.42 (p .08). The difference between the conventional and validation groups was also not significant. This non-significant difference between these latter two groups was nevertheless meaningful since the conventional method employed instructions requesting that the individual react to the items in a personal way, and the validation group instructions emphasized socially desirable responding. Moreover, these statistical findings directly implied that socially acceptable modes of responding are intricately woven into one's personal belief system, making it difficult to separate one from the other.

Dissection of the treatment conditions on the basis of internality and externality affords still another opportunity to call on \( t \)-tests to examine and compare cell means (see Table 20 for between cell \( t \)-tests). For example, comparing internals to externals in the bogus pipeline treatment revealed a very strong statistically significant relationship. Similar findings occurred for the internal and external groups embedded in the conventional situation. The validation subgroup, on the other
hand, was non-significant which suggested that internals' scores were not easily distinguishable or different from externals' scores. In fact, this can be clearly seen by inspecting the corresponding means for the internal and external cells in the validation condition. In addition, the validation category was the only group of the entire sample in which the mean for externality was less than the mean for internality.

Checking cell means across treatment conditions was also undertaken. The mean score of internals in the bogus pipeline method was statistically different from the mean of internals functioning under conventional procedures ($t=1.94$, $p=.05$). A test of the means of conventional and validation internals produced a $t=-3.16$, $p=.001$. However, the examination of the bogus pipeline and validation means for internals did not furnish a meaningful difference.

In a similar fashion the means of the external cells were paired across treatment situations and analyzed. The analysis of the means of externals in the bogus pipeline and conventional methods indicated that the difference was at best only minimal. Inspection of the means for externals in the conventional and validation conditions by way of the $t$-test showed that these cell means to be in sharp contrast. This was also characteristic of the relationship between the bogus pipeline and the validation means for externals.

A complete means breakdown of the experimental design can be located in the appendix (see Table 14). A breakdown is merely collapsing over one or more of the independent factors with the purpose of offering dif-
ferent viewpoints on the behavior of the criterion variable (i.e. dependent variable). The mean value of the criterion variable in addition to the corresponding cell size and standard deviation are listed.

The following sections will discuss at length the meaning of this study's findings and their implications.
CHAPTER IV
DISCUSSION

On the whole, the results of this investigation are consistent throughout. Admittedly, the issue of the effects of social desirability on the locus of control's outcome has already be widely researched (Cone, 1971; Hjelle, 1971; Rotter, 1966; Vuchinich and Bass, 1974), but the current study offers at the very least a different method for documenting this influence. Although the focus of the study was to examine social desirability in responding to the items on the I-E scale, reliability estimates and item analysis of the scales were also conducted. The purpose of computing the reliability estimates and undertaking an examination of the items of the scales was twofold. The internal consistency check and the item analysis data replicated findings reported elsewhere (Rotter, 1966). In addition, these procedures were used to estimate the extent of compatibility among scale items for the sample population. Understandably, knowing whether the test items of the I-E scale and SD scale are measuring a single construct respectively or several constructs was needed before any firm conclusions could be drawn. This was particularly true of the I-E scale whose dimensionality is currently being contested in some research quarters. Researchers (Abramowitz, 1973; Collins, 1973; Levenson, 1973 and others) raising doubt over the scale's dimensionality contend that the test items are not sampling the same underlying attitude. Therefore, they believe the scale to be multidimensional in nature. The other group of researchers,
led by the work of Rotter, have shown with some regularity that the scale is unidimensional, that it is measuring a specific construct. Moreover, the bulk of research suggests that the scale is factorially homogeneous. The study's replication of Rotter's results provided support for the unidimensional explanation of the scale's internal structure. The .72 reliability coefficient found for the initial I-E scale testing was respectable when examined with respect to the scale's format. The forced-choice format of the scale, though a helpful means of lessening response bias, tends also to influence the computed reliability coefficient. In other words, the lack of independence among forced-choice type of items usually tends to lower the internal consistency estimate (Kerlinger, 1973).

Interpreting the internal consistency of the I-E test items at the treatment level becomes a more complex enterprise since the subject's interaction with the items was being purposely manipulated either, directly or indirectly. The decrease in the internal consistency estimate for subjects who experienced the bogus pipeline situation during the second testing of the I-E scale signals a slight alteration in the items' capacity to act as indicators of a particular underlying attitude. Furthermore, it is well-known that forced-choice scales can strain the subject's patience, resulting in less cooperation (Kerlinger, 1973). In the bogus pipeline situation respondents were compelled to chose one of the item's alternatives before 15 seconds elapsed. This might well have further eroded the patience of subjects and thus caused the mild drop in the reliability estimate. Still it must be remembered that the 15 seconds
interval allotted after the presentation of the alternatives of an item was adequate because it gave the subject ample time to reflect and then to decide. Consequently, it was felt that the institution of a reaction time was consistent with the notion of forced-choice since both typify the nature of daily, routine decision-making activity of human beings.

The second administration of the I-E scale to the conventional and validation groups produced increases in reliability estimates. The large coefficient recorded for the conventional sample was expected. The conventional condition, the control group of the study, was designed to replicate the first I-E examination. Therefore, the relatively slight increase in the coefficient is acceptable, even predictable given the familiarity of the subjects to the scale items.

The fairly high reliability estimate found for the second I-E presentation for the validation group was not unreasonable, nor was it unexpected. Subjects in this case were told to make their selection contingent upon what they believed to be the most socially desirable of the two alternatives for each item. Instructing subjects to respond in a way which they considered to be most desirable by society would tend to create a uniform response pattern amongst them. As illustrated in the present case, internal consistency coefficient was increased.

Several item analysis procedure were selected to examine how well items of the same scale measured the same construct. The specific inter-correlational procedures were item-to-item, uncorrected item-to-total, and corrected item-to-total correlations. The item-to-total correlational matrix for both the first I-E and second I-E testing
appeared to be free from any major suspicious trend. The item-to-
total correlation technique is perhaps the most straightforward item
analysis procedure (Bohrnstedt, 1970). In this technique one simply
takes each item and correlate it with the total scale score. The uncor-
corrected and the corrected item-to-total correlations for the I-E scale in
both phases of the study were reasonable and tended to support the uni-
dimensionality of the scale. In fact, the corrected item-to-total cor-
relations obtained for the first I-E administration, though only moderate,
closely approximate those correlations presented by Rotter (1966) for a
combined male and female population. The item analysis data gathered on
the SD scale also seemed relatively free of gross correlational dis-
crepancies.

Reliability estimates and item analysis are indeed requisite pro-
cedures to employ in a study in which scales are involved. The informa-
tion obtained from these techniques can provide an index for judging the
scales' stability. Also, the interpretation of the data gathered on the
scale's internal structure provides confidence in the research. The
worth of reliability procedures is obvious. Yet reliability checks can
only examine certain aspects of a scale's construction. The impact of
social desirability on the scale's items would be difficult to assess
with reliability procedures, unless it was embedded in the scale itself.

This study, however, was particularly interested in trying to docu-
ment the effect of social desirability on the I-E scale's outcomes. The
experimental situations of the project were constructed so that social
desirability was being experimentally controlled. The bogus pipeline
treatment, for example, was designed to reduce social desirability bias on the locus of control scale. The primary purpose of the validation condition was to heighten responses tainted with social desirability and to clarify the direction of the socially attractive responding to the I-E scale items. Of course, the conventional procedure, a traditional paper and pencil method, represented the usual mode of administering the locus of control scale. It was thus hypothesized that these three experimental situations would differ markedly from each other. To test this main effect as well as other main effects and interactional hypotheses an analysis of variance approach was chosen.

The statistically significant F-test computed for the treatment conditions suggests at the very least that the experimental manipulation accorded these three groups was appreciably different. An examination of the treatment means showed that the bogus pipeline, the conventional, and the validation groups recorded means as anticipated. That is, the group mean for the bogus pipeline sample was the most external, while the validation group registered the lowest mean score indicating the most internally-oriented responses. The mean score for conventional subjects fell between the other two experimental groups. The direction and the distribution of the treatment means bear upon several points of the project's purpose. First, when subjects were persuaded that their real attitudes were being electronically measured they tended to become more externally minded. Subjects, confronted with an elaborate-looking apparatus usually found themselves mildly apprehensive and anxious. However, this may facilitate a truer response by the subject because the individual may be
fearful of being out-smarted and embarrassed by a machine. In other words, the subject's need to depict himself in a socially attractive manner may give way to his need to avoid being contradicted by and therefore "found out" by a machine. In essence, subjects are more likely to be open about their real feelings when they are made to think that those feelings can be extracted independent of their conscious control. The validation group freed subjects to respond without any reservations in a socially desirable way. The general finding was that the social attractive alternatives on the I-E scale were predominately the internal choices. A group of pairwise $t$-tests further illuminated the relationship between the treatment conditions. This statistic revealed that the bogus pipeline subject population was different from the other two treatment samples, even though only the $t$-test comparing the bogus pipeline and the validation treatment means was significant at the .05 alpha level. Evaluation of the means for the conventional and validation situations gave clear indication that the difference between these two different conditions was small. Moreover, the inability of the validation and the conventional groups to be differentiated from each other may bring into question the utility of the I-E scale. In particular, whether one's generalized expectancy is internal or perhaps even external may be contingent upon what the subject believes to be the most socially appealing, rather than what he/she feels to be the truest alternative for him/her.

The significant main effect found for the first I-E testing was fully expected. In the first I-E presentation no experimental influence
was employed, therefore the scores were normally distributed. However, the dependent variable which entailed the second I-E testing at the treatment levels was greatly affected by the experimenters' manipulations. Consequently, the comparison of the first I-E outcomes, a set of normally distributed scores with the dependent measure which had been experimentally exposed resulted in the statistically meaningful F-test. In the identification stage, the initial testing phase of the project, subjects were given the I-E and the SD scales. After this stage the scores from both scales functioned as independent variables. The introduction of treatment conditions in the second phase changed the role of the I-E scale. Depending upon the situation the I-E scale was exposed to either direct, or indirect, or no experimental influence. The objective of this aspect of the study was to determine if the manipulation of the social desirability factor would induce modifications in I-E outcomes. Examining the study's overall findings it can be concluded that the employment of the I-E scale as an independent and dependent variable did not result in any confounding effects.

Testing the main effect for both the SD scale and sex proved non-significant. However, there was great interest in whether sex influenced outcomes on the dependent measure. In fact, the heightened interest stemmed from the fact that several I-E researchers have reported conflicting findings on this topic. Specifically, Feather (1968) has claimed that the locus of control score for females tend to be more external than males. The opposing position held by Rotter (1966), Gore (1963) and others contend that the difference between sexes on the locus of control scale is minuscule and statistically non-significant. In the present study sex
did not have a distinguishable or unique relation with the dependent variable. Inspection of Figure 1 does show that females are more likely to earn higher external as well as higher internal scores; the bar graph presentation further demonstrates the differences to be non-appreciable.

The first order interaction of the treatment conditions with the first I-E presentation was highly significant. This interaction, like the main effect found for the first I-E testing, was logical. Similarly, subjects in the beginning phase of the project received the same instructions as to how to approach the scale's items whereas subjects in the treatment situations were given different instructions, according to the experimental group. This variation in how subjects perceived the items on the second locus of control scale examination lead to the highly significant finding. The second order interaction of the treatment groups with the first I-E administration and sex was statistically significant at the .05 level. This finding suggested that sex when considered simultaneously with the treatment groups and the I-E test was a distinguishing factor.

The other specific aim of the study, though not analyzed by the multivariate procedure, contends that the bogus pipeline's internals are more likely to have scores leaning towards externality than conventional internals. The most optimal test to use to examine this relationship was the t-test. This statistical method supported the hypothesis that bogus pipeline internals who were made to believe that their true attitude across an array of different situations could be successfully monitored differed markedly from internals assigned to the traditional pencil and paper setting.
Subjects with internal scores appear to be more likely than externals to be affected by social desirability. In other words, internals appear to have the most difficulty in distinguishing their own feelings from what is socially desirable. Hjelle (1971) showed that a sizeable number of internal items were significantly more socially desirable than the corresponding external items. Similarly, Cone (1971) suggested that internals who feel that they have some control over the type of reinforcement received are more inclined to behave in socially desirable ways.

Vuchinich and Bass (1974) concluded that highly internally-oriented individuals tend to have high needs for social approval. This however does not seem to be the case for externals. Externals in the present study came across more in-touch with their feelings since very little shifting across treatment conditions was observed. Furthermore, externals were less likely than their counterparts to be swayed by what was socially approved and sanctioned. Comparison of the means for the bogus pipeline and the conventional externals also attest to this fact. The non-significant t value of .23 (p .41) implied that externals regardless of whether they are experimentally manipulated or not will tend to respond more externally and are steadfast in their responses to the I-E scale.

This study has shown that social desirability responding can be experimentally modified under certain circumstances. Social desirability in the bogus pipeline situation appear to have been lessened. Subject's fear of being "found out" by a machine controlled to some extend response bias. Asking subjects in the validation condition to respond to the I-E scale in the most socially desirable way indicated that internality was
the most socially attractive direction. Interesting to note is that internal males in the validation situation leaned heavily toward the external aspect of the scale. This may be due to the fact that males in this society are usually taught that they are in control of their environment. Thus the relinquishing of such control may be construed as a socially undesirable act by them. In addition to documenting the effects of social desirability across treatment settings, the project has demonstrated that subjects with internal scores are perhaps more homogeneous group than externals because internals react to the I-E scale in a socially desirable manner. The following section will examine the implications of the study's findings.

IMPLICATIONS

The study's findings imply that the I-E scale when employed as a pencil and paper task may have difficulty tapping a subjects' true feelings or beliefs. A major factor which this study hypothesized to be contributing to the scale's difficulty was social desirability. That is, subjects when responding to the items on the scale might often fuse their personal feelings with what was for them the most socially acceptable alternatives. This fusion of one's personal feelings with social desirability could result in distorted or unrepresentative impressions of subjects' generalized expectancies. To get a closer look at the effects of social desirability, the study's treatment conditions were constructed to manipulate this factor. It was shown that the bogus pipeline paradigm, a system designed to elicit a true response, was less
likely to be contaminated with socially desirable type of responding. On the other hand, the conventional and validation groups were both more likely to have scores inflated due to social desirability; subjects in these two conditions, for instance, were instructed to respond to the locus of control items in a personal way and in a socially attractive manner, respectively. Although these groups received contrasting instructions, there was no difference in their overall response patterns. This suggests that subjects in the conventional situation tended to choose the more socially attractive item alternatives. It also appeared to indicate that personal beliefs and perceived social pressure are intricately fused together and can not be completely differentiated. This seemed to be characteristic of internals more than externals. As mentioned, internals have been described as representing a fairly homogeneous grouping (Hersch and Scheibe, 1967); whereas externals are considered to be a more amorphous grouping. Nevertheless, the results of this study suggest that the apparently tight, uniform internal grouping might be due to subjects' tendency to respond to the items of the I-E scale in a socially desirable fashion. The validation treatment condition which aided in the inspection and illumination of the extent of influence of social desirability in subjects' response patterns, showed that socially desirable responding on the I-E scale was likely to be in the internal direction. In fact, this condition was the most internally-directed of the three experimental groups. Moreover, it seemed that the more external the sample population the smaller the affect of treatment manipulation on I-E outcomes. A comparison of bogus pipeline system internals
revealed that under the bogus pipeline system internals were markedly more externally-oriented than internals in the conventional setting. The monitoring of a subject's reactions affect how he will respond to the I-E items and as suggested by the post-questionnaire his responses may be more closely related to his personal feelings under such conditions. On the other hand, subjects in the conventional method were not monitored and were consequently more inclined to adopt society's views, even though their personal opinions were requested.

A difference was also found on the post-questionnaire between internals and externals assigned to the bogus pipeline method. Internals, for instance, endorsed the system's usefulness more frequently; whereas externals as indicated on the post-questionnaire tended to be more cautious, even suspicious.

One of the more striking findings of the project showed that externals in the bogus pipeline and conventional situation were not markedly different from each other. The absence of any significant difference between these two groups implies that externals are more likely to express their personal feelings. This means that externals, unlike internals, are less likely to respond to the I-E scale items on the basis of whether the responses are socially desirable. Perhaps this explains at least in part why externals tend to represent a heterogeneous grouping. That is, external subjects' personal reactions were seemingly more likely to be offered, while internals as illustrated by the study's findings appeared over-concern with being viewed as socially attractive.

In addition to demonstrating the affects of social desirability on I-E outcomes, it was also noted that the mean score of 11.0 with a
standard deviation of 4.82 on the first I-E testing was considerably higher than the 8.05 average score with a 3.74 standard deviation reported by Rotter (1966) for combined sample population of males and females. The variability of I-E outcomes over the years can be construed as an indication of the influence of social desirability. For instance, the present study was conducted during the midst of the Watergate scandal and the winding down of United States' involvement in Southeast Asia which may explicate the higher external mean score for the overall sample population. These two events were significant in creating a state of disillusionment among a majority of the people regarding the present social and political systems. Although subjects' personalities may not have been changed, their views about the social and political issues may have been modified in an external direction by such events. In this case the I-E scale may not have been tapping aspects of personal control. Or as Coan (cf. Dies, 1968) puts it, the I-E scale favors items dealing with social and political events as opposed to items regarding personal habits, traits, goals, or other interpersonal and intrapersonal concerns.

This study was also interested in investigating the utility of the bogus pipeline procedure. Earlier reports by Sigall and Jones (1971) and Sigall and Pages (1971) have been highly favorable concerning the power of the bogus pipeline technique. The results of the present study lends some additional support to this system of deception for the purpose of eliciting a more accurate reflection of a subject's feelings, beliefs, or attitudes. However, like most systems of deception, the bogus pipeline does have some drawbacks. The bogus pipeline, though more accurate than the conventional pencil and paper method, is time consuming and a highly
masked procedure which may not be useful for certain kinds of experiments, especially if the experiments are unable to effectively add the parameter of deception into the experimental design. Moreover, deception techniques on the whole have a relatively short life span. The increased number of subjects being exposed to the bogus pipeline type of deception plus the reporting of this and similar procedures in such widely read magazines as the Psychology Today and the New York Times will eventually rendered this technique useless. Obviously, the excessive use of deception has basically made the naive subject a near extinct resource for psychological research. Nevertheless, the bogus pipeline system offers at least for the time being a mode of examining subject's personal feelings.

Finally, the argument of the study focussed not on the I-E construct but on the locus of control scale devised by Rotter to measure it. Specifically, the study was interested in determining if a certain dimension of the subject's personality was being measured as suggested by the scale's objective or if the scale was incline to assess the subject's view of what was desirable and approved by social system. The study showed that social desirability influenced significantly I-E outcomes. This was particularly evident for internals whose need for social approval often made them lose sight of their personal feelings. However, because this project viewed internals differently than most of the I-E literature, it would be important to replicate the study's procedures as well as undertake other means of documenting the effects of social desirability in subjects' responses to the locus of control scale. This investigation has
also shown that current reliability measurements though highly useful are insufficient and must be supplemented with other modes of determining if a scale is truly an indicator of a specific psychological construct.
SUMMARY

The investigation evaluated the impact of social desirability on I-E outcomes and found it to be a salient factor in how subjects respond to the locus of control scale. 247 male and female subjects were initially given the I-E and SD scales. After a 15 day time interval subjects were randomly assigned to one of the three treatment situations --the bogus pipeline, the conventional, and the validation--and readministered the I-E scale. Subjects placed in the bogus pipeline condition were persuaded that their true feelings were being electronically monitored, thereby making it difficult for them not to respond to the items on the scale in a personal way. The conventional situation constituted for the most part the control condition of the experiment since no experimental manipulation was used. The validation condition, like the conventional situation, was a pencil and paper procedure. However, subjects in the validation situation were instructed to respond to the I-E items in a socially desirable manner. An analysis of variance procedure revealed that the treatment conditions were significantly different. The bogus pipeline sample group was the most externally-oriented while the group mean for the validation sample was the most internally-directed. In addition, conventional internals differed markedly from bogus pipeline internals, whereas the bogus pipeline and the conventional externals were not significantly different. It was also found that internals in the bogus pipeline situation were far more likely to endorse this method of securing someone's true attitude than externals. In sum, the results support the study's hypothesis that I-E outcomes may be distorted by social desirability responding. This was especially true for internals
in the conventional condition who displayed a greater tendency to respond in a socially sanctioned and approved fashion.
REFERENCES


Goss, A., and Morosko, T.E. Relation between a dimension of internal-ex-


Hersch, P.D., and Scheibe, K.E. On the reliability and validity of internal-external control as a personality dimension. Journal of Con-
Hjelle, L.A. Social desirability as a variable in the locus of control scale. Psychological Reports, 1971, 28, 807-816.


Strickland, B.R. Locus of control: where have we been and where are we going? Paper presented at the American Psychological Association, 1973, Montreal, Canada.


APPENDICES
APPENDIX A: ROTTER'S I-E SCALE

1. a. Children get into trouble because their parents punish them too much.
   b. The trouble with most children nowadays is that their parents are too easy with them.

2. a. Many of the unhappy things in people's lives are partly due to bad luck.
   b. People's misfortunes result from the mistakes they make.

3. a. One of the major reasons why we have wars is because people don't take enough interest in politics.
   b. There will always be wars, no matter how hard people try to prevent them.

4. a. In the long run people get the respect they deserve in this world.
   b. Unfortunately, an individual's worth often passes unrecognized no matter how hard he tries.

5. a. The idea that teachers are unfair to students is nonsense.
   b. Most students don't realize the extent to which their grades are influenced by accidental happenings.

6. a. Without the right breaks one cannot be an effective leader.
   b. Capable people who fail to become leaders have not taken advantage of their opportunities.

7. a. No matter how hard you try some people just don't like you.
   b. People who can't get others to like them don't understand how to get along with others.

8. a. Heredity plays the major role in determining one's personality.
   b. It is one's experiences in life which determines what they're like.

9. a. I have often found that what is going to happen will happen.
   b. Trusting to fate has never turned out well for me as making a decision to take a definite course of action.

10. a. In the case of the well prepared student there is rarely if ever such a thing as an unfair test.
    b. Many times exam questions tend to be so unrelated to course work that studying is really useless.

11. a. Becoming a success is a matter of hard work, luck has little or nothing to do with it.
    b. Getting a good job depends mainly on being in the right place at the right time.
12. a. The average citizen can have an influence in government decisions.
   b. This world is run by few people in power, and there is not much the little guy can do about it.

13. a. When I make plans, I am almost certain that I can make them work.
   b. It is not always wise to plan too far ahead because many things turn out to be a matter of good or bad fortune anyhow.

14. a. There are certain people who are just no good.
   b. There is some good in everybody.

15. a. In my case getting what I want has little or nothing to do with luck.
   b. Many times we might just as well decide what to do by flipping a coin.

16. a. Who gets to be the boss often depends on who was lucky enough to be in the right place first.
   b. Getting people to do the right thing depends upon ability, luck has little or nothing to do with it.

17. a. As far as world affairs are concerned, most of us are the victims of forces we can neither understand, nor control.
   b. By taking an active part in political and social affairs the people can control world events.

18. a. Most people don't realize the extent to which their lives are controlled by accidental happenings.
   b. There really is no such thing as "luck."

19. a. One should always be willing to admit mistakes.
   b. It is usually best to cover up one's mistakes.

20. a. It is hard to know whether or not a person really likes you.
   b. How many friends you have depends upon how nice a person you are.

21. a. In the long run the bad things that happen to us are balanced by the good ones.
   b. Most misfortunes are the result of lack of ability, ignorance, laziness, or all three.

22. a. With enough effort we can wipe out political corruption.
   b. It is difficult for people to have much control over the things politicians do in office.

23. a. Sometimes I can't understand how teachers arrive at the grades they give.
   b. There is a direct connection between how hard I study and the grades I get.
24. a. A good leader expects people to decide for themselves what they should do.
b. A good leader makes it clear to everybody what their jobs are.

25. a. Many times I feel that I have little influence over the things that happen to me.
b. It is impossible for me to believe that chance or luck plays an important role in my life.

26. a. People are lonely because they don't try to be friendly.
b. There's not much use in trying too hard to please people, if they like you, they like you.

27. a. There is too much emphasis on athletics in high school.
b. Team sports are an excellent way to build character.

28. a. What happens to me is my own doing.
b. Sometimes I feel that I don't have enough control over the direction my life is taking.

29. a. Most of the time I can't understand why politicians behave the way they do.
b. In the long run the people are responsible for bad government on a national as well as on a local level.

Note---items with neither alternative underlined are filler items.
Conventional Method: Social Reaction Inventory

This is a questionnaire to find out the way in which certain important events in our society affect different people. Each item consists of a pair of alternatives lettered a or b. Please select the one statement of each pair (and only one) which you more strongly believe to be the case as far as you're concerned. Be sure to select the one you actually believe to be more true rather than the one you think you should choose or the one you would like to be true. This is a measure of personal belief: obviously there are no right or wrong answers.

Please answer these items carefully but do not spend too much time on any one item. Be sure to find an answer for every choice. Find the number of the item on the answer sheet and black-in the space under the number 1 or 2 which you choose as the statement more true.

In some instances you may discover that you believe both statements or neither one. In such cases, be sure to select the one you more strongly believe to be the case as far as you're concerned. Also try to respond to each item independently when making your choice; do not be influenced by your previous choices.

Validation Method

This is a questionnaire to find out which of the two alternatives for each item is the most socially acceptable.

Please read the alternative for each item carefully and then choose the one that seems to be the most approved by society. Remember! It is important that you make your selections on what you think is socially acceptable since our own viewpoints on specific topics often times differ from the viewpoint that is most socially acceptable.
APPENDIX C: PERSONAL REACTION INVENTORY
(MARLOWE-CROWNE SOCIAL DESIRABILITY SCALE)

Listed below are a number of statements concerning personal attitudes and traits. Read each item and decide whether the statement is true or false as it pertains to you personally.

Do not make any marks on the test booklet. Record your answers in the true or false columns of the separate answer sheet that has been given you. Fill in your name and sex on the answer sheet.

Remember: Answer each item as it pertains to you personally.

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

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

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

4. I have never intensely disliked anyone. +

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

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

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

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

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

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

11. I like to gossip at times. -

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

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

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

15. There have been occasions when I took advantage of someone. -
16. I'm always willing to admit it when I make a mistake. +
17. I always try to practice what I preach. +
18. I don't find it particularly difficult to get along with loud
mouthed obnoxious people.
19. I sometimes try to get even rather than forgive and forget. -
20. When I don't know something I don't at all mind admitting it. +
21. I am always courteous, even to people who are disagreeable. +
22. At times I have really insisted on having things my own way. -
23. There have been occasions when I felt like slamming things. -
24. I would never think of letting someone else be punished for my
wrongdoings. +
25. I never resent being asked to return a favor. +
26. I have never been irked when people expressed ideas very different
from my own. +
27. I never make a long trip without checking the safety of my car. +
28. There have been times when I was quite jealous of the good fortune
of others. -
29. I have almost never felt the urge to tell someone off. +
30. I am sometimes irritated by people who ask favors of me. -
31. I have never felt that I was punished without cause. +
32. I sometimes think when people have a misfortune they only got what
they deserved. -
33. I have never deliberately said something that hurt someone's feelings. +
APPENDIX D: SUBJECT'S INTRODUCTION TO EMG

(Pre and Post Questionnaires)

This equipment is called an "adapted electromyograph." The EMG, its abbreviated name, is a device used to measure "implicit muscle potentials." The EMG has been heralded by psychological researchers as an important and new break-through which circumvents many of the methodological problems associated with standard questionnaires.

The EMG apparatus employs four electrodes. Two of the four electrodes are placed on the forearms; and the other two electrodes are attached to the palm of the left hand. These electrodes function to allow the EMG to screen-out gross-muscle movements and to record the first undistorted reaction. In essence, the equipment you're viewing works as a system to integrate the electro-physiological input and this (pointing to it), the EMG output meter, will reflect a close approximation of that integration. Furthermore, it is important to remember that this machinery is not a lie detector. In fact, the EMG is an improvement over the lie detector in that it is sensitive to direction as well as intensity of responses. Therefore, you will notice that the meter pens have been removed and that those signals usually read by the meter pens (pointing to them) will be incorporated into the EMG system.

---ATTACH SKIN ELECTRODES TO SUBJECT---

Occasionally base-line responses differ among individuals, in which case an adjustment of the EMG is required. To check this we simply go over the items on the pre-questionnaire and compare your responses to the EMG readings. Now, please listen carefully to the following statements and after each statement look, if you wish, at the EMG output reading. You'll notice that a reading going to the right will indicate a disagreement with the statement while the opposite direction signals agreement. To put this in somewhat different terms, no verbal response is necessary from you to measure your true reaction since the EMG records a reflection of your first, undistorted reaction. During the course of the calibration you might try fooling the EMG either by exercising gross muscle movements or consciously concentrating on the direction opposite your true position. You will find upon comparison that such strategies are not successful.

---COMPARE EMG REACTIONS TO SUBJECT'S PRE-QUESTIONNAIRE RESPONSES---

The purpose of the present experiment is to examine personal perceptions. Although this information can be obtained directly from the EMG, we are also interested in how sensitive you are to your own feelings. That is, whether you are in touch with how you feel. Therefore, without looking at the EMG output readings we would like you to predict what you think the meter will say.
PRE - QUESTIONNAIRE

Read each item carefully and decide whether the statement is true or false for you. Please circle your answer.

1. The most important thing for a parent to do is to help (his/her) children get further ahead in the world than (he/she) did.
   - True
   - False

2. I am an only child.
   - True
   - False

3. The average man is probably better off today than he ever was.
   - True
   - False

4. I currently smoke cigarettes.
   - True
   - False
POST - QUESTIONNAIRE

Please answer the following questions. Try to be as clear and specific as possible.

1. Briefly state your impressions upon entering the experimental room.

____________________________________________________________________

____________________________________________________________________

____________________________________________________________________

____________________________________________________________________

2. Please describe what you believe to be the purpose of the study.

____________________________________________________________________

____________________________________________________________________

____________________________________________________________________

____________________________________________________________________

3. Is this procedure an effective way of securing a person's true attitude? YES _____ NO _____ (please check) -------
   Why or why not: EXPLAIN.

____________________________________________________________________

____________________________________________________________________

____________________________________________________________________

____________________________________________________________________

4. Please use the rest of the form for further elaboration or additional comments.

____________________________________________________________________

____________________________________________________________________

____________________________________________________________________

____________________________________________________________________

Thank you!
Dear Participant:

Upon signing up for this experiment, you were given an initial explanation which described the purpose of this study as an investigation of the formation of attitudes. Indeed, this project was interested in examining the attitudes and beliefs of its respondents. However, the examination of attitudinal variables was not the primary purpose; instead the central aim of the study was to ascertain whether responses to the Social Reaction Inventory (i.e. Rotter's Internal-External Locus of Control Scale) are influenced by the desire to be seen as socially acceptable. All subjects used in this experiment were randomly assigned to one of the three experimental conditions. The conventional condition merely requires each subject to darken in on an optic scan sheet those alternatives on the I-E scale that characterized their feelings and attitudes. The conventional condition represents the control for this experiment. Subjects placed in the validation and EMG Conditions experienced some experimental manipulation. Specifically, subjects in the validation condition were instructed to respond to Rotter's 29 item forced-choice scale in the "most socially acceptable way." The third condition, the EMG, is based on the lie-detector premise that one cannot fool electronic devices which are geared to measure physiological changes. In truth the apparatus used in the experiment was nothing but electronic junk. Its abilities (e.g meter readings) were regulated by a confederate (an assistant) who was housed in the adjacent room and had access to your preliminary questionnaire. The confederate's specific function was to manipulate the meter pen so that it corresponds to your preliminary questionnaire. This constitutes the major effort to persuade the subjects in this condition that deliberately false responses were immediately detectable. Nevertheless, the professed EMG equipment used in the EMG condition is nothing more than an elaborate system of deception and cannot measure "implicit muscle potential".

Often psychological research has found itself somewhat handicapped when studying why people respond to items on a questionnaire in a certain way. In particular, psychology has had difficulty finding direct, straightforward but still effective means of obtaining the desired data from subjects. Because of this difficulty psychology, especially in experimental situations with human subjects, has frequently employed deceptive methods for securing information pertinent to the problem under investigation.
Nevertheless, we regret having to utilize deception but understandably, it was impossible to disclose the real nature of the study without ruining its principle purpose.

Furthermore since deception is incorporated into the experimental design of this study, we would appreciate it if you would refrain from discussing the experiment with friends and other students. However, please feel free at this time to use the back of this page for comments and suggestions. You can forward your responses to:

William G. Harris
Tobin Hall
Psychology Department
University of Massachusetts
Amherst, Massachusetts  01002

Finally, your participation in this experiment was gratefully appreciated. Thank you!

Sincerely,

William G. Harris
Experimenter
APPENDIX F: PERSONAL DATA SHEET

NAME: 

IDENTIFICATION: 

ADDRESS: 

PHONE NUMBER: 

MAJOR: UNDECLARED 

AGE: SEX: RACE: 

YEAR IN SCHOOL: 

OVERALL GRADE POINT AVERAGE: 

NUMBER OF PSYCHOLOGY COURSES COMPLETED: 

GRADE POINT AVERAGE IN PSYCHOLOGY COURSES: 

RESIDENTIAL BACKGROUND: 

1,000-25,000

25,000-100,000

100,000+

PREDOMINANTLY RURAL:

PREDOMINANTLY SMALL TOWN:

PREDOMINANTLY TOWN:

PREDOMINANTLY CITY:
APPENDIX G: TABLES AND FIGURE
# Table 1

**Internal and External Frequency Tabulations of Subject's Reactions as Related to High and Low Need for Approval**

<table>
<thead>
<tr>
<th>I-E/SD Groups</th>
<th>SEX</th>
<th>Approve</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Frequency</td>
<td>Percent</td>
<td>Frequency</td>
</tr>
<tr>
<td><strong>INT. HIGH</strong></td>
<td>FEMALES</td>
<td>9</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>MALES</td>
<td>9</td>
<td>90</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>18</td>
<td>95</td>
<td>1</td>
</tr>
<tr>
<td><strong>INT. LOW</strong></td>
<td>FEMALES</td>
<td>7</td>
<td>70</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>MALES</td>
<td>4</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>11</td>
<td>61</td>
<td>4</td>
</tr>
<tr>
<td><strong>EXT. HIGH</strong></td>
<td>FEMALES</td>
<td>5</td>
<td>63</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>MALES</td>
<td>5</td>
<td>63</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>10</td>
<td>63</td>
<td>5</td>
</tr>
<tr>
<td><strong>EXT. LOW</strong></td>
<td>FEMALES</td>
<td>5</td>
<td>56</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>MALES</td>
<td>4</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>9</td>
<td>53</td>
<td>4</td>
</tr>
</tbody>
</table>
### Table 2
Internal, External, and Overall Frequency Tabulations of Subject’s Reactions to the Bogus Pipeline Situation

<table>
<thead>
<tr>
<th>GROUPS</th>
<th>SEX</th>
<th>APPROVE</th>
<th></th>
<th>DISAPPROVE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Frequency</td>
<td>Percent</td>
<td>Frequency</td>
<td>Percent</td>
</tr>
<tr>
<td>OVERALL</td>
<td>FEMALES</td>
<td>26</td>
<td>72.2</td>
<td>7</td>
<td>19.4</td>
</tr>
<tr>
<td></td>
<td>MALES</td>
<td>22</td>
<td>65.0</td>
<td>7</td>
<td>21.0</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>48</td>
<td>69.0</td>
<td>14</td>
<td>20.0</td>
</tr>
<tr>
<td></td>
<td>FEMALES</td>
<td>16</td>
<td>84.0</td>
<td>2</td>
<td>11.0</td>
</tr>
<tr>
<td></td>
<td>MALES</td>
<td>13</td>
<td>72.0</td>
<td>3</td>
<td>17.0</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>29</td>
<td>78.0</td>
<td>5</td>
<td>14.0</td>
</tr>
<tr>
<td></td>
<td>FEMALES</td>
<td>10</td>
<td>59.0</td>
<td>5</td>
<td>29.0</td>
</tr>
<tr>
<td></td>
<td>MALES</td>
<td>9</td>
<td>56.0</td>
<td>4</td>
<td>25.0</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>19</td>
<td>58.0</td>
<td>9</td>
<td>27.0</td>
</tr>
</tbody>
</table>
The alpha reliabilities for the I-E and SD scales were .70 and .75, respectively.

<table>
<thead>
<tr>
<th>I-E</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>33</td>
<td>63</td>
</tr>
<tr>
<td>34</td>
<td>64</td>
</tr>
<tr>
<td>35</td>
<td>65</td>
</tr>
<tr>
<td>36</td>
<td>66</td>
</tr>
<tr>
<td>37</td>
<td>67</td>
</tr>
<tr>
<td>38</td>
<td>68</td>
</tr>
<tr>
<td>39</td>
<td>69</td>
</tr>
<tr>
<td>40</td>
<td>70</td>
</tr>
<tr>
<td>41</td>
<td>71</td>
</tr>
<tr>
<td>42</td>
<td>72</td>
</tr>
<tr>
<td>43</td>
<td>73</td>
</tr>
<tr>
<td>44</td>
<td>74</td>
</tr>
<tr>
<td>45</td>
<td>75</td>
</tr>
<tr>
<td>46</td>
<td>76</td>
</tr>
<tr>
<td>47</td>
<td>77</td>
</tr>
<tr>
<td>48</td>
<td>78</td>
</tr>
<tr>
<td>49</td>
<td>79</td>
</tr>
<tr>
<td>50</td>
<td>80</td>
</tr>
</tbody>
</table>

Initial testing for the bogus procedure sample.

Inter-item correlations and corrected and uncorrected item-total score correlations for the I-E scale.
The alpha reliabilities for the 10 and 12 scales were .73 and .69, respectively.

**Table 1**

<table>
<thead>
<tr>
<th>Initial Testing for the Conventional Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item-Item Correlations and Corrected Unrevised Item-Total Score Correlations for the I-E Scale</td>
</tr>
</tbody>
</table>
The alpha reliabilities for the I-R and SB scales were .64 and .70, respectively.

<table>
<thead>
<tr>
<th>I-R</th>
<th>SB</th>
</tr>
</thead>
<tbody>
<tr>
<td>.64</td>
<td>.70</td>
</tr>
</tbody>
</table>

Table 9

Inter-item correlations and corrected inter-item-total score correlations for the I-R scale.
TABLE 11

The Correlation of Need for Approval with both the Initial and Experimental I-E Scale Testings

<table>
<thead>
<tr>
<th>I-E</th>
<th>Social Desirability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bogus Pipeline</td>
</tr>
<tr>
<td>I-E (1)</td>
<td>-.16</td>
</tr>
<tr>
<td>I-E (2)</td>
<td>-.26*</td>
</tr>
</tbody>
</table>

*p .05

1 The overall correlation found for the initial I-E (1) and SD testing was -.15, p .01
TABLE 12

Test-retest Reliability for Treatment Conditions

<table>
<thead>
<tr>
<th></th>
<th>Bogus Pipeline</th>
<th>Convention</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-E (1)</td>
<td>.79</td>
<td>.85</td>
<td>-.11</td>
</tr>
<tr>
<td>I-E (2)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 13
Analysis of Variance: of Treatment Groups for the Experimental Testing of the I-E Scale

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Groups (A)</td>
<td>2</td>
<td>62.33</td>
<td>3.85*</td>
</tr>
<tr>
<td>I-E (1) (B)</td>
<td>1</td>
<td>853.28</td>
<td>52.77***</td>
</tr>
<tr>
<td>SD (C)</td>
<td>1</td>
<td>.99</td>
<td>.06</td>
</tr>
<tr>
<td>Sex (D)</td>
<td>1</td>
<td>.04</td>
<td>.00</td>
</tr>
<tr>
<td>A X B</td>
<td>2</td>
<td>365.64</td>
<td>22.61***</td>
</tr>
<tr>
<td>A X C</td>
<td>2</td>
<td>42.83</td>
<td>2.65</td>
</tr>
<tr>
<td>A X D</td>
<td>2</td>
<td>13.82</td>
<td>.85</td>
</tr>
<tr>
<td>B X C</td>
<td>1</td>
<td>.42</td>
<td>.03</td>
</tr>
<tr>
<td>B X D</td>
<td>1</td>
<td>15.37</td>
<td>.95</td>
</tr>
<tr>
<td>C X D</td>
<td>1</td>
<td>47.44</td>
<td>2.93</td>
</tr>
<tr>
<td>A X B X C</td>
<td>2</td>
<td>.28</td>
<td>.02</td>
</tr>
<tr>
<td>A X B X D</td>
<td>2</td>
<td>49.95</td>
<td>3.09*</td>
</tr>
<tr>
<td>A X C X D</td>
<td>2</td>
<td>7.92</td>
<td>.49</td>
</tr>
<tr>
<td>B X C X D</td>
<td>1</td>
<td>8.07</td>
<td>.50</td>
</tr>
<tr>
<td>A X B X C X D</td>
<td>2</td>
<td>2.43</td>
<td>.15</td>
</tr>
<tr>
<td>S/ABCD</td>
<td>223</td>
<td>16.17</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>246</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p .05
***p .001
TABLE 15

Means, Standard Deviations, and Sample Size on the Experimental I-E Scale Testing in terms of I-E (1), Sex, and Treatment Group

<table>
<thead>
<tr>
<th>I-E (1)</th>
<th>Sex</th>
<th>Mean</th>
<th>Std.</th>
<th>n</th>
<th>Mean</th>
<th>Std.</th>
<th>n</th>
<th>Mean</th>
<th>Std.</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bogus Pipeline</td>
<td>Conventional</td>
<td>Validation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internals</td>
<td>F</td>
<td>32.25</td>
<td>3.61</td>
<td>20</td>
<td>31.52</td>
<td>3.78</td>
<td>23</td>
<td>32.00</td>
<td>5.20</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>32.00</td>
<td>3.87</td>
<td>17</td>
<td>30.67</td>
<td>3.84</td>
<td>26</td>
<td>35.00</td>
<td>3.83</td>
<td>26</td>
</tr>
<tr>
<td>Externals</td>
<td>F</td>
<td>38.35</td>
<td>3.84</td>
<td>17</td>
<td>37.94</td>
<td>2.98</td>
<td>22</td>
<td>33.64</td>
<td>5.12</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>37.81</td>
<td>2.40</td>
<td>16</td>
<td>37.91</td>
<td>2.80</td>
<td>19</td>
<td>32.25</td>
<td>5.48</td>
<td>20</td>
</tr>
</tbody>
</table>
### TABLE 16

Means, Standard Deviations, Sample Sizes on the Experimental I-E Scale Testing in terms of I-E (1)

<table>
<thead>
<tr>
<th>I-E (1)</th>
<th>Bogus Pipeline</th>
<th>Conventional</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Means</td>
<td>Std.</td>
<td>n</td>
</tr>
<tr>
<td>Internals</td>
<td>32.68</td>
<td>3.74</td>
<td>37</td>
</tr>
<tr>
<td>Externals</td>
<td>38.09</td>
<td>3.19</td>
<td>33</td>
</tr>
<tr>
<td>Treatment Conditions</td>
<td>Mean</td>
<td>Std.</td>
<td>n</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------</td>
<td>-------</td>
<td>-----</td>
</tr>
<tr>
<td>Bogus Pipeline</td>
<td>35.23</td>
<td>4.40</td>
<td>70</td>
</tr>
<tr>
<td>Conventional</td>
<td>34.19</td>
<td>4.82</td>
<td>90</td>
</tr>
<tr>
<td>Validation</td>
<td>33.44</td>
<td>4.94</td>
<td>87</td>
</tr>
<tr>
<td>Total</td>
<td>34.22</td>
<td>4.78</td>
<td>247</td>
</tr>
</tbody>
</table>
TABLE 18

Means, Standard Deviations and Sample Sizes on the Experimental I-E Scale Testing in terms of the Sex Variable

<table>
<thead>
<tr>
<th>Sex</th>
<th>Mean</th>
<th>Std.</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Females</td>
<td>34.39</td>
<td>4.86</td>
<td>123</td>
</tr>
<tr>
<td>Males</td>
<td>34.06</td>
<td>4.72</td>
<td>124</td>
</tr>
</tbody>
</table>
TABLE 19

Between Treatment Condition T-Tests for the Experimental I-E Scale Testing

<table>
<thead>
<tr>
<th>Condition</th>
<th>Conventional</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bogus Pipeline</td>
<td>1.42</td>
<td>2.40**</td>
</tr>
<tr>
<td>Conventional</td>
<td>1.03</td>
<td></td>
</tr>
</tbody>
</table>

**p .01
one-tailed
TABLE 20

Between Cell T-Tests for the Experimental I-E Scale Testing

<table>
<thead>
<tr>
<th>Treatment/I-E Groupings</th>
<th>Pipe+E</th>
<th>Conv-I</th>
<th>Conv-E</th>
<th>Valid-I</th>
<th>Valid-E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pipe-I</td>
<td>-6.55***</td>
<td>1.94*</td>
<td>-6.91***</td>
<td>-1.29</td>
<td>-.35</td>
</tr>
<tr>
<td>Pipe-E</td>
<td>-9.05***</td>
<td>.23</td>
<td>-4.68***</td>
<td>5.28***</td>
<td></td>
</tr>
<tr>
<td>Conv-I</td>
<td>-9.76***</td>
<td>-3.16***</td>
<td>-2.05*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conv-E</td>
<td>-4.83</td>
<td>5.43***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valid-I</td>
<td></td>
<td></td>
<td></td>
<td>.82</td>
<td></td>
</tr>
</tbody>
</table>

*p .05  
***p .001

1Separate variance estimates were used to test the relationships between cells
Figure 1. A bar graph of the mean scores for treatment groups on the first I-E testings as related to the subjects' sex.

<table>
<thead>
<tr>
<th>BOSS PIPETTE</th>
<th>CONVENTIONAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- M: Male
- F: Female

I-E Mean Scores

10
20
30
40