CALIFORNIA STATE UNIVERSITY, NORTHRIDGE

JUSTIFICATION OF DECEPTION FROM THE
SUBJECT’S POINT OF VIEW

A thesis submitted in partial satisfaction of the requirements for the degree of Master of Arts in Psychology

by

James William Myers

June, 1976
The thesis of James William Myers is approved:

Dr. Paul Skolnick (Date)

Dr. Susan Shodahl (Date)

Dr. James McMartin (Date)  Committee Chairman

California State University, Northridge
ACKNOWLEDGEMENT

I wish to thank the members of my committee for their help and understanding during the writing of this paper. To Dr. James McMartin for his time, thoughts and valuable suggestions, to Dr. Susan Shodahl for her enthusiastic support, and to Dr. Paul Skolnick for coming in on his off-time. I also wish to thank Cindy whose patience and encouragement made this thesis possible.
### TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vii</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>METHOD</td>
<td>22</td>
</tr>
<tr>
<td>Subjects</td>
<td>22</td>
</tr>
<tr>
<td>Procedure</td>
<td>22</td>
</tr>
<tr>
<td>Personal Consequence Manipulation</td>
<td>22</td>
</tr>
<tr>
<td>Scientific Importance Manipulation</td>
<td>25</td>
</tr>
<tr>
<td>Experimenter Concern Manipulation</td>
<td>27</td>
</tr>
<tr>
<td>RESULTS AND DISCUSSION</td>
<td>29</td>
</tr>
<tr>
<td>MANIPULATION CHECKS</td>
<td>29</td>
</tr>
<tr>
<td>Personal Consequence</td>
<td>29</td>
</tr>
<tr>
<td>Scientific Importance</td>
<td>30</td>
</tr>
<tr>
<td>Experimenter Concern</td>
<td>33</td>
</tr>
<tr>
<td>DEPENDENT MEASURES</td>
<td>34</td>
</tr>
<tr>
<td>Justifiability of Deception</td>
<td>34</td>
</tr>
<tr>
<td>Perception of Deception in Terms of Ethics</td>
<td>37</td>
</tr>
<tr>
<td>Subjects Evaluation of Experiment</td>
<td>38</td>
</tr>
<tr>
<td>Willingness to Participate</td>
<td>38</td>
</tr>
<tr>
<td>Importance of Experimentation in Terms of Usefulness of Results</td>
<td>39</td>
</tr>
<tr>
<td>Importance of Experimentation in Terms of the Welfare of Subjects</td>
<td>42</td>
</tr>
<tr>
<td>CONCLUSION</td>
<td>43</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>47</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>50</td>
</tr>
<tr>
<td>A. COVER SHEET</td>
<td>51</td>
</tr>
<tr>
<td>B. INTRODUCTION TO AX-TYPE EXPERIMENT WITH DR. HOWARD BRADFORD</td>
<td>53</td>
</tr>
<tr>
<td>C. INTRODUCTION TO AX-TYPE EXPERIMENT WITH HOWARD BRADFORD</td>
<td>54</td>
</tr>
<tr>
<td>D. INTRODUCTION TO MILGRAM-TYPE EXPERIMENT WITH DR. HOWARD BRADFORD</td>
<td>55</td>
</tr>
<tr>
<td>E. INTRODUCTION TO MILGRAM-TYPE EXPERIMENT WITH HOWARD BRADFORD</td>
<td>56</td>
</tr>
<tr>
<td>F. INTRODUCTION TO FREEDMAN AND DOOB-TYPE EXPERIMENT WITH DR. HOWARD BRADFORD</td>
<td>57</td>
</tr>
<tr>
<td>G. INTRODUCTION TO FREEDMAN AND DOOB-TYPE EXPERIMENT WITH HOWARD BRADFORD</td>
<td>58</td>
</tr>
<tr>
<td>H. INTRODUCTION TO ARONSON AND LINDER-TYPE EXPERIMENT WITH DR. HOWARD BRADFORD</td>
<td>59</td>
</tr>
<tr>
<td>I. INTRODUCTION TO ARONSON AND LINDER-TYPE EXPERIMENT WITH HOWARD BRADFORD</td>
<td>60</td>
</tr>
<tr>
<td>J. AX-TYPE EXPERIMENT</td>
<td>61</td>
</tr>
<tr>
<td>K. MILGRAM-TYPE EXPERIMENT</td>
<td>63</td>
</tr>
<tr>
<td>L. FREEDMAN AND DOOB-TYPE EXPERIMENT</td>
<td>65</td>
</tr>
<tr>
<td>M. ARONSON AND LINDER-TYPE EXPERIMENT</td>
<td>66</td>
</tr>
<tr>
<td>N. QUESTIONNAIRE</td>
<td>67</td>
</tr>
</tbody>
</table>
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Cell means for Scientific Importance X Debriefing in response to &quot;How justifiable was the deception in terms of possible benefit to mankind?&quot;</td>
<td>36</td>
</tr>
<tr>
<td>2.</td>
<td>Cell means for Stress X Debriefing in response to the statement &quot;The most important aspect of an experiment is how useful are the results.&quot;</td>
<td>40</td>
</tr>
<tr>
<td>3.</td>
<td>Cell means for Scientific Importance X Debriefing in response to the statement &quot;The most important aspect of an experiment is how useful are the results.&quot;</td>
<td>41</td>
</tr>
</tbody>
</table>
ABSTRACT

JUSTIFICATION OF DECEPTION FROM THE SUBJECT'S POINT OF VIEW

by

James William Myers

Master of Arts in Psychology

The ethics of experimentation with human subjects is a topic of continuing controversy. Justification for possibly unethical procedure such as deception, is found by experimenters in subjects' reports that the experiment was a worthwhile experience and that they were glad to have participated. This thesis attempted to illustrate that the subject's report may be strongly influenced by factors other than the actual procedures used. It was hypothesized that the scientific importance of the study, the concern of the experimenter, and the amount of stress involved are three such factors.

Tested in their classrooms, 240 male and female introductory psychology students at California State
University, Northridge, received booklets containing descriptions of an experiment involving deception. This description also included a short narrative regarding the experimenter varying his prestige and goals of research. An indication of the extent of the debriefing was also present.

A 3 x 2 x 4 design (debriefing x scientific importance x stress) was employed to test the hypotheses that subjects would report deceptive procedures as more justified when they perceived the stimulus experiment to be high in scientific importance and when they perceived the stimulus experimenter to be very concerned about his subjects. Additionally it was believed that these effects would occur only in stressful stimulus experiments.

Main effects for all three variables were obtained supporting the first two but not the third hypothesis. The absence of an interaction between stress and scientific importance or debriefing indicates that subjects are strongly affected by these factors regardless of the amount of stress involved.

Since it has been shown that the experimenter can elicit varying degrees of justifiability from a subject depending on the alleged value of his undertaking and the display of concern during the debriefing, care
must be taken not to excuse possibly unethical procedures merely because the subjects have said it is alright in the end.
"Scientific ends, however laudable these may be, do not by themselves justify the use of means that in ordinary transactions would be regarded as reprehensible."

Baumrind (1971)

The question of the ethics of scientific research often revolves around this "end justifies the means" concept. The common response is that if we believe the acquisition of scientific knowledge about human behavior is a positive value, and if an experiment using questionable methods is considered a significant contribution to such knowledge which could not be achieved by other means, then we cannot absolutely rule out this experiment. The sacrifice, or more accurately, discomfort, of a few individuals is a fair price to pay for the advancement of knowledge and betterment of society in general. Or at least this is what some scientists would have us believe.

One of the difficulties with this way of thinking is that the determination of the extent to which the risk to the individual benefits society is purely subjective. Although there may be an implied semi-mathematical relationship in the phrase "the end justifies the means"
[(end ≥ means) = justification], an attempt to quantify either "end" or "means" would be to lessen the inherent meaning of both. If these concepts were operationally defined it is highly likely that the definitions would be as varied in nature as are the views expressed by each side of the controversy.

In a reexamination of the American Psychological Association's Code of Ethics, Cook, Kimble, Hicks, Schoggen, and Smith (1971) agreed on certain ethical principles regarding the ends/means proposition. Basically these principles offer the investigator appealing rationalizations for risking the subject's welfare based upon the anticipated benefit to society. This permits him to expose an uninformed subject to possible psychological harm provided he has sufficient reason, an advisory committee behind him, and is willing to try to undo whatever harm the subject experiences (Baumrind, 1971).

One of the "means" that has come under ethical fire is the use of deception in experimental procedures. The dilemma arises as the researcher realizes that his significant area of experimentation cannot be investigated adequately without deceiving his subjects in one way or another and a conflict of values ensues as he must decide if his ends are equal to or greater than his
means. Kelman (1967) states that awareness of this dilemma is at least a partial solution, since if the investigator is sufficiently sensitive to the possible psychological consequences of deception he may actively guard against unwittingly inflicting harm. The use of deception then becomes a matter for discussion, deliberation, investigation, and choice. Alternative procedures may be available and would be ethically preferable to the use of deception. There may be instances, however, where alternative procedures are impractical. Aronson and Carlsmith (1968), along with several other investigators, believe that one of the best methods of obtaining experimental realism is through the use of deception. This is not accomplished, though, without some cost, usually to the unsuspecting subject.

Two recurring themes in the case against deception are 1) the influence upon the subject as he views himself and 2) the impact on his relationships with others. Danger to the subject's self-image is expressed by Walster, Berschield, Abrahams, and Aronson (1967) in a study assessing the effectiveness of the debriefing procedure. They claim that some deceptive manipulations are emotionally disturbing to a subject and that some disturbances might not be entirely relieved by the post-experimental debriefing. John Jung (1971) agrees that
deception is a source of harm in experiments in which the subject is made to feel inferior or foolish even when he is told later that he had been misled into forming these self-impressions.

Kelman (1967) addresses the issue of the impact on the subject's social relationships when he asserts that deception does not merely take place within the experiment, but encompasses the entire definition of the relationship between the parties involved. The subject learns, after being repeatedly lied to by highly respected scientists, that he cannot trust those whom he needs to trust in order not to feel alienated by society (Baumrind, 1964; Walster et al., 1967). These notions of alienation and distrust of people in general are as yet intuitive speculation with little, if any, empirical foundation.

One further concern is voiced by Ring who fears for the respectability of the profession in the eyes of fledgling psychology students. He asks "what is the perspective student to think of a field where the most renowned researchers apparently get their kicks from practicing sometimes unnecessary and frequently crass deceptions on the unsuspecting subjects?" (1967, p. 118). He does not, however, condone the abolition of all deception; only unnecessary deception. In reference to
Milgram's (1963) study on obedience, Ring acknowledges that although it raised some serious ethical questions it was a good example of both interesting and significant research.

It is clearly an unending debate with neither side the victor and empirical evidence hard to obtain due to the very nature of the issue. Those who would argue against deception would not be comfortable using it as a manipulation in an experiment and therefore would not obtain data. Those who argue for deception rarely see the need to justify its use.

A common suggestion regarding alternatives is that researchers should invest some of the creativity and ingenuity now devoted to the designing of elaborate deceptions to the search for experimental techniques that do not rely on deception. Kelman (1967) calls for a method that would elicit the subject's positive motivations to contribute to the experimental enterprise and draw upon the subject's active participation and involvement.

One suggested alternative to deception is the use of naturalistic studies. Baumrind (1964) believes that to the extent an area of research is socially significant, it will exist naturally in the real world. The researcher can then use these naturally occurring
phenomena as an alternative to laboratory studies where deception must be used. Although this would seem to be a proper solution for eliminating the deception of subjects, it gives rise to other ethical objections. Ruebhausen and Brim (1965) claim that field observations in spite of the success of "Candid Camera" - involve one of the most heinous ethical violations in our culture: invasion of privacy.

Ethical problems are also evident in field studies in which investigators pose as members of a special interest group so they can observe its operation from the inside. The most notable example is the investigation of the "doomsday group" observed by Festinger, Riecken, and Schachter (1956). The ethical dilemma of naturalistic studies is further acknowledged by Webb, Campbell, Schwartz, and Sechrest (1966) who suggest further debate on the matter.

In addition to the ethical objections, naturalistic studies also generate many methodological problems. Strict controls are impossible to achieve; samples may be unrepresentative of the population depending on time of day, day of week, etc.; and often one has to wait for naturally occurring phenomena to occur naturally.

Another frequently proposed alternative to deception is role playing. Unfortunately from a
methodological viewpoint, role playing is not a totally satisfactory remedy. Aronson and Carlsmith (1968) maintain role playing lacks the realism required to secure generalizable results. Miller (1972) points out that the difficulties in role playing are similar in nature to the problems arising from introspective reports. In addition, the role player has at least three potential preoccupations: 1) the behavior he would be likely to perform if the situation were real, 2) the behavior he should perform to be in line with society's norms, and 3) the behavior he thinks the experimenter expects of him. The last two of these objections may be used to benefit rather than hinder an experiment.

To illustrate, Orne (1969) suggested that role playing be used as a quasi-control procedure to assess the effects of demand characteristics. In this context it would not be an alternative to deception, but rather a method to detect how much of a given experimental finding is due to the subject's ability to see through the deceptions and artifactually provide support for the experimenter's hypothesis. These demand characteristics are defined as the totality of cues to which the subject responds that influence his behavior in the experiment (Orne, 1962). Subjects need not even be aware that they have been cued and, in fact, when asked if there had been any influence on their behavior, most reply there had not. Of course there is no guarantee that this
Orne (1959) recognized this possibility and described what he terms the pack of ignorance. The subject who has seen through the deceptions and is aware of what is really happening may hesitate to admit it because he realizes he may be disqualified from the experiment. This consideration is undesirable for the experimenter as well and he may refrain from probing too deeply in the post-experimental inquiry. Adair (1973), on the other hand, warns of just the opposite. Too many leading questions in the post-experimental inquiry may suggest to the subject that the experimenter really wants him to be aware or will think he is stupid. This may be further complicated by the fact that the role attitude that governed the subject's behavior during the experiment may still be operating during the inquiry. These role attitudes may take the form of cooperation or defensiveness.

Several investigators (Baumrind, 1964, 1971; Adair, 1973; Orne, 1962) have referred to an implicit personal contract between the subject and the experimenter which is sealed at the time a subject volunteers or signs up. By doing so the subject "agrees" to assume a posture of trust and obedience. In return he has the right to expect the psychologist with whom he is interacting to have some concern for his welfare and the personal attributes and professional skills to express his good will effectively (Baumrind, 1964). If these
conditions appear to be fulfilled the result is the cooperative role attitude of the subject.

However, if the experiment is designed to evaluate the subject personally, i.e., measures of his abilities, capabilities, attitudes, or personal traits, he may develop a defensive or apprehensive attitude (Adair, 1973). Obtaining the subject's reactions to aggression, dependency, altruism, or pressures to conform, sometimes causes a certain degree of anxiety and stimulates efforts to look good. Weber and Cook (1972) demonstrated that evaluation apprehension is probably the major source of subject-caused artifacts. This apprehension may exist by itself or in interaction with other subject roles.

It is true that the alternatives to deception are as yet insufficiently developed and refined. Each has its own problems and, to reiterate the tagline of most psychological studies, further research is required.

The published research aimed at comparing deception with its alternatives center mostly around the methodological problems. Admittedly this form of inquiry is essential to uphold the validity of experimentation, but an irony exists in the paucity of experiments regarding the effects on the subject. It seems most experimenters are so engrossed in their quest for alternatives to deception or support for deception they overlook the object of their concern - the subject. The subject bears the weight and consequences of unethical methods.
Aside from Ring (1967) who believes deception is part of the "fun and games" psychologists play, most of us would probably agree that well trained experimenters deceive subjects only because they want to create experimental realism and not because they enjoy fooling or hurting people. In fact, most experimenters attempt to uncover or tap into the subject's real feelings about being deceived by the use of careful post-experimental debriefing. Often during an interview, or as an item on a questionnaire, the subject will be asked in one way or another if he thought his deception was justified. Intuitively one might suspect that in general people do not like to be deceived. This is alluded to in several ways by many researchers. Aronson and Carlsmith (1968) contend that most people do not enjoy learning that they are gullible and have been taken in. Baumrind (1964) reports how one subject felt humiliated and embarrassed by a laboratory ruse. Campbell (1969), in referring to the debriefing process, states that it may be painful for subjects to learn of their gullibility, conformity, cruelty, or bias. This "inflicted insight" into one's character may be harmful indeed. Baumrind (1964) declares that she would expect a naive, sensitive subject to remain deeply hurt and anxious for some time and a sophisticated cynical subject to become even more alienated and distrustful.

In light of these sentiments from the
profession's finest, it is difficult to accept the reports of so many subjects who found their own deception justified. For instance, two studies by Bramel (1962, 1966) involved leading young students to believe they had homosexual tendencies. Kelman (1967) contends that for many persons of this age group sexual identity is still a live and sensitive issue and the self doubts caused by the laboratory experience may take on a life of their own and linger on for some time. Yet, according to Bramel "all available evidence indicates that the subjects considered the experience interesting and worth their while." (1966, p. 320).

Aronson and Linder (1965) manipulated self-esteem by giving some of their female subjects negative evaluations stating they were dull conversationalists, rather ordinary, not very intelligent, and probably did not have many friends. After debriefing many subjects were relieved to learn it was all a hoax. Aronson and Linder report "although several of the girls admitted to having been quite shaken during the experiment, they felt it was a worthwhile experience inasmuch as they learned the extent to which a negative evaluation can affect them. ... They left the interview room in good spirits" (p. 162). Here we have a good example of the notion of inflicted insight.

Rosenthal and Jacobson (1968) led teachers to believe some of their students (selected at random) were
"intellectual bloomers." Consequently these students, over a period of time, increased their IQ scores significantly more than the nonintellectual bloomers. When the study was complete and the deceptions explained, none of the teachers expressed resentment or thought the deception was unethical. In fact, the authors reported a sense of excitement in the teachers for having played a part in what they regarded to be important research. One wonders how these teachers felt about the randomly selected (by default) non-intellectual bloomers whose IQ scores might have been higher had they been the lucky ones.

Berkun, Bialek, Kern, and Yagi (1962) devised several experiments on stress that convinced Army recruits that their lives were truly in danger. In one such experiment, an isolated subject in a desolate area learned that a sudden emergency had arisen (accidental nuclear radiation, a sudden forest fire, or misdirected artillery fire - depending on the experimental condition) and the only means of rescue was through his radio transmitter which just happened to malfunction at that time. During the debriefing session the subject was given an opportunity to vent any residual tensions and negative feelings. One or two weeks later a second interview was conducted and in the words of Berkin, et al.
"in no case was there any evidence of residual tension or negative feelings" (p. 12). How easily they forget!

The preceding experiments make one wonder if Kelman's (1967) general principle of treatment is being followed. He proclaims that the subject ought not to leave the laboratory with greater anxiety or lower self-esteem than he came in with. The defenders of the deception technique have pointed out that subjects in nearly all deception studies are thoroughly debriefed at the end of the experiment, and this debriefing will accomplish the goal of returning the subject to his pre-experimental state. For example, Holmes and Bennet (1974) found debriefing to be effective for reducing the level of stress in experimental subjects to the level of non-stressed control subjects. The question arises, nevertheless, how can the researcher be absolutely sure that his debriefing has been successful in this way. Walster et al. (1967) found debriefing to be ineffective even after a long time delay, for certain personality types. Ross, Lepper, and Hubbard (1975) provide convincing evidence that there is a definite perseverance of self-perceptions based on erroneous information when a standard debriefing is performed.

It would be unfortunate indeed if, as Campbell (1969) distressingly suggests, the debriefing is more a
comfort to the experimenter to assuage his guilt over
deceiving subjects than it is to the subject who is
likely to be more in need of that comfort. It has also
been charged that to allay possible ethical objections,
many experimenters inform their readers that "standard"
reassurance is supplied at the end of the experiment,
with the implication that subjects go away happy and
proud about their "guinea pig" role (Vinacke, 1954).

Currently there does not seem to be a workable
solution to the dilemma. Even if the experimenter is
sincere in his attempts to reconcile any harm done, the
subject may find it difficult to express his anger and
hostility outwardly when the experimenter, in a friendly
and kind manner, reveals the hoax. Remaining silent
about his objections, then, the subject leads the exper-
imenter to believe he has performed a successful debrief-
ing.

One partial solution is offered by Kelman (1967)
although it violates certain tenets of randomization and
subsequent generalizability. He suggested that subjects
be selected so that those who are especially vulnerable
to psychological harm will be excluded. Assuming that
we actually had knowledge to make this selection ac-
curately, we would ameliorate the problems pointed to by
Walster et al. (1967).
Two further experiments employing deception are appropriate for discussion regarding the interaction between subject and experimenter in the debriefing process. Freedman and Doob (1968) were interested in deviancy. The basic procedure in a series of experiments was to give the subject five personality tests and then inform him that the results indicated he was "quite different from most people." In the debriefing, considerable time was spent explaining that the feedback was erroneous. As an added precaution, the authors went over it again and again until it was certain that everyone understood the deceptions. They noted that in many cases the subject required some convincing since apparently the manipulation was both believable and impressive and the subjects did in fact feel deviant. In spite of this, however, Freedman and Doob insisted that since the feedback lacked content, it did not bother them as much as a seemingly less powerful manipulation might have and that the debriefing appeared to be sufficient to remove whatever effect the manipulation had.

Two very important considerations are overlooked by these experimenters. First, it seems that having the deception explained in detail over and over again would certainly be embarrassing to the subject and become more so with every recounting of his gullibility. Thus a
subject may feign satisfaction and relief just to be able to exit himself from the ever increasingly humiliating situation. Secondly, the belief that contentless deviancy is not very harmful is to ignore some of the more basic notions of psychology. Since none of us is perfect, the knowledge acquired from the bogus personality tests that we are deviant in some respect from the rest of society may direct our attention to that one secret self-doubt we've had for some time, but managed to ignore or suppress for lack of any concrete evidence. But alas, here is that evidence we were afraid of! It was true all along that we were really (choose one) conformists, sadists, dullards, stupid, obedient, effeminate, masculine, cowardly, or just plain weird. Each subject now has the evidence that his particular self-doubt is authentic. How could this be any less harmful; it excludes no one! What does it really mean when Freedman and Doob report they were satisfied that the debriefing was totally successful?

A final and most controversial illustration is offered by Milgram (1963). This well-known study of obedience has been the focus of severe criticisms in the issue of ethics in research. Baumrind (1964), in perhaps the most scathing review, maintains it is potentially harmful to a subject to commit, in the course of an
experiment, acts which he himself considers wrong or unethical, especially when he has been trapped into doing so by someone he has reason to trust. The stark realization that he is capable of hurting others in such a manner could easily effect an alteration of the subject's self-image and his willingness to trust adult authority in the future. An additional lowering of self-esteem could occur merely from the subject's belief that he made a fool of himself by accepting the experimental set.

In response to such objections Milgram (1964) insists that procedures were employed to assure that the subject left the experiment in a state of well being and that every effort was made to reduce any tension that arose during the experiment. Milgram's description of the extreme tension that existed included nervous laughter fits, sweating, trembling, stuttering, groaning, digging one's fingernails into the flesh, and in three cases full blown uncontrollable seizures.

It is not doubted that Milgram took numerous steps to thoroughly debrief and restore his subjects to normality. He saw that as his duty as a scientist and as a feeling individual. Yet in spite of his good intentions there is sufficient reason to believe that at least some of the obedient subjects sustained a lower self-esteem, having to live with the fact that they
yielded to destructive authority to the point of
inflicting extreme injury on another human being. The
fact that they "learned something of importance about
themselves" is beside the point.

In further defense of his experiments, Milgram
maintains that neither he nor his colleagues could have
foreseen how far the subjects would actually go. The
question then became one of continuing such research or
abandoning it because of the amount of stress experi­
enced by the subjects. A quote from Milgram's (1964)
reply to Baumrind (1964) serves as kindling for the
hypotheses of the present paper:

It is true that after a reasonable number of
subjects had been exposed to the procedures,
it became evident that some would go to the
end of the shock board, and some would experi­
ence stress. That point, it seems to me, is
the first legitimate juncture at which one
could even start to wonder whether or not to
abandon the study . . . and as subjects them­
selves strongly endorsed the experiment, the
judgement I made was to continue the investi­
gation. (p. 849)

Milgram and other researchers appear to use as
justification for deception and other questionable
methods, the subject's "endorsement" of these techniques.
It is proposed that when a subject, during some post-
experimental inquiry, states or implies that his own
deception was justified, he is succumbing to the demand
characteristics of the situation. He may, in fact,
honestly believe his deception was not justified. This would seem to be an important distinction for those interested in the ethics of experimentation, yet very little research has been conducted to assess the subject's true feelings taking into account the possible contamination of demand characteristics.

This contamination may develop from the cooperative role attitude described by Fillenbaum (1966) who declares that insofar as the experimenter can demonstrate that a deception had a legitimate and necessary purpose for the sake of the experiment, the subject will dutifully go along, trusting and helping him in his research. Expanding on this, one can assume that the subject will additionally consider the importance of the experiment in terms of scientific importance or social applicability. The promise of science to remedy our problems and give us a better life has led to a passive acceptance of various policies and procedures that result in presumed benefit to the majority at the expense of the minority. At the individual level, one is convinced that his relatively small amount of suffering is inconsequential when viewed in the light of possible good that may come from it. More to the purpose of this paper, if a subject in an experiment believes that his suffering or embarrassment will result in the alleviation
of some social malady, he will be more likely to brave the deception and discomforts that occur and report these occurrences as justified. The demands upon him to do so are quite apparent. After all, which one of us would refuse to endure a small amount of uneasiness if many others less fortunate than ourselves would benefit in the long run?

It is also possible that the subject may be reacting to the experimenter's concern for his welfare as established by the quality of the debriefing. The more involved the debriefing, the greater the presumed concern of the experimenter. Even if a subject believes that the deceptions he went through are not justifiable, he may be reluctant to express his anger to the apologetic and "well meaning" experimenter.

However, it seems plausible that only when a deception leads to high personal consequences (i.e., those involving stress for the subject) will experimenter concern and scientific importance have an effect on the subject's report of the justifiability of the deception. For low consequence deceptions (no stress), there should be very little discomfort for the subject and the question of deception justifiability becomes academic.

The purpose of this thesis is to attempt to separate the influence of the presumed concern of the
experimenter from the demand characteristics of scientific importance or social applicability. A 3 x 2 x 4 design (debriefing x scientific importance x stress) will be employed to test the following hypotheses: 1) subject will rate a deception higher in justifiability when the outcome of the experiment contributes to the understanding and resolution of some social problem than when the outcome has no obvious social applicability; 2) subjects will rate a deception higher in justifiability when, during the debriefing, the experimenter shows concern for the subject's welfare than when the experimenter shows little or no concern; and 3) the effects of scientific importance and experimenter concern will have an effect only when the deception is one that places the subject under some form of stress.¹

¹Although the primary concern of this investigation is the subject's own reaction to possibly unethical procedures, subjects will not actually experience these procedures. Rather, they will read about the experiences of another. The reports of observers cannot be assumed to be identical to those of actual participants and, in fact, may be more conservative since there is lower personal consequences for observers.
METHOD

Subjects

All subjects were undergraduate students enrolled in introductory psychology courses at California State University at Northridge, who participate for course credit.

Procedure

Two hundred and forty subjects were randomly assigned to booklets, the contents of which were a function of the experimental condition. All instructions were given in the booklets.

The first page of each booklet contained an introduction to the study and a statement of its purpose; i.e., an examination of the effects of certain experimental procedures on subjects. In order to encourage true and accurate responses, subjects were reminded that they could very well participate in experiments of the type they read about while fulfilling their requirements for experimental credits.

Personal Consequence Manipulation

Ethics are often seen on a continuum with some actions being more or less ethical than others. Deceptions which in no way arouse emotional conflicts,
self-doubts, or anxiety may well be deemed justified even by some of the opponents of deception. For example, the communicator credibility studies of Hovland, Janis, and Kelley (1953) which deceived subjects into believing a communication came from either a high or low credible source could hardly be accused of inflicting emotional or psychological harm. On the other end of the continuum, the Milgram (1963, 1965) experiments on obedience and the Berkun et al. (1962), studies on stress might be seen as objectionable and unjustified even to some of the proponents of deception.

In order to identify both ethically objectionable (stressful) and non-objectionable (non-stressful) experiments from the viewpoint of the subjects, descriptions of ten experiments involving deception were presented to thirty-eight subjects from introductory psychology classes. Each subject read three of the ten studies. After each description, subjects indicated, among other things, whether or not they thought the deceptions employed were justified in relation to the costs incurred by the subjects and in relation to the potential gain in psychological knowledge. From these, two experiments were chosen such that the majority of the subjects deemed the deception unethical, and two were chosen such that the majority deemed the deception not unethical. Half
the subjects in the main experiment read the stressful studies, half read the non-stressful studies.

The two stressful experiments selected were as follows:

Ax (1953). A variation of Ax's study led subjects to believe that equipment had short-circuited and there existed a danger of being electrocuted.

Milgram (1963). This famous, or infamous, study of obedience has received severe criticism on ethical grounds for inducing subjects to believe they were inflicting serious harm on another.

The two non-stressful experiments chosen were as follows:

Freedman and Doob (1968). An adaptation of this study of deviancy was employed using a milder, less negative description of the subject's departure from normality.

Aronson and Linder (1965). Subjects were exposed to the "gain" condition of this gain/loss treatment of attraction. A staged favorable conversation about a subject was overhead by him during the course of the experiment.
Scientific Importance Manipulation

As yet, there has not been much investigation of this variable and a simple "high-low" manipulation of one factor would probably be insufficient. There are numerous contributing variables that affect one's interpretation of scientific importance and it is not known exactly which of these are the most influential. Since this investigation is not designed to identify the major components of that which labels an endeavor scientifically important, several factors thought to contribute to this notion were combined at each extreme, resulting in two levels which intuitively suggest high and low scientific importance.

These factors included: 1) the contribution of the outcome of the experiment to the resolution of some social problem, 2) the prestige of the experimenter, 3) the competence of the experimenter, 4) the sponsor and funding of the research, and 5) the location of the experiment.

An example of the combination of these factors for high scientific importance is as follows: Dr. Bradford, a well known psychologist working at Harvard University under National Science Foundation Grant Number 1056, was conducting experiments on the psychological effects of
stress on affiliation. The outcome of this particular experiment, in all likelihood, would culminate eight years of intensive research and supply Dr. Bradford with a major breakthrough in answering the questions of affiliative behavior. With the knowledge gained from this research it will be possible for people of different cultures, races, and beliefs, to live in more harmony than has ever been experienced in the past. As always, Dr. Bradford was very efficient in the execution of his experiments and was able to make optimum use of his time and resources. He used subjects drawn from introductory psychology courses, as he did in all previous experiments, and proceeded as follows ...

An example of the combination for low scientific importance is:

Howard Bradford, a senior at Goshen College in Indiana, working with funds from the Student Research Association, was conducting experiments on the psychological effects of stress on affiliation. The outcome of this particular experiment will, in all likelihood, not contribute very much to the understanding of affiliative behavior in
terms of harmony between different cultures and races, but will satisfy the requirements of one of Howard's class assignments. As always, Howard was not very efficient in the execution of his experiments and seldom made optimum use of his time and resources. He used subjects drawn from introductory psychology courses, as he did in all previous experiments, and proceeded as follows . . .

Experimenter Concern Manipulation

Three levels of debriefing were described in the booklets and were crossed with each of the above variables. Subjects in the No Debriefing condition found the following statement at the end of their experiment: "At the conclusion of this experiment, subjects were thanked for their time but were not informed of the true nature of the deceptive procedures nor the real purpose of the study." Subjects in the Short Debriefing condition read: "At the conclusion of this experiment Dr. Bradford (or Howard) thanked the subjects for their time and spent about ten minutes describing the deceptive procedures and the major purpose of the study. He also answered questions the subjects had." The Thorough Debriefing condition contained the following message: "At the conclusion of this experiment, Dr. Bradford (or
Howard) thanked the subjects for their time and spent about thirty minutes explaining, in detail, the deceptive procedures and the major purpose of the study. In addition to answering all questions subjects might have had, he probed to see if any after effects of the deception persisted. If he found this to be so, he continued to discuss the deception with the subjects until he was absolutely sure that all subjects left the experiment in the same psychological condition they came in with."
RESULTS AND DISCUSSION

MANIPULATION CHECKS

Personal Consequence

Subjects rated the experiment they read about on a seven point scale ranging from "harmful" (1) to "harmless" (7) which served as a check on the personal consequence or stress manipulation. It is reasonable to assume that a situation without stress could not very easily be termed harmful while one involving stress might definitely be viewed as harmful. An F ratio of 14.51 (df = 3,216, p<.001) was obtained and a Tukey test for multiple comparisons revealed both stressful experiments were significantly more harmful than both non-stressful experiments.

Additionally, the perception of harm was affected by the scientific importance variable. Experiments presented as being low in scientific importance were rated as more harmful than those high in scientific importance: F (1,216) = 4.16, p<.05. This lends support to the notion that if a subject believes his suffering will benefit others he is less likely to report his discomfort.

A main effect of experimenter concern was also evident in the ratings of harmfulness of the experiment:
A Tukey test comparing the three levels of debriefing indicated that a short debriefing ($\bar{X} = 3.25$) was significantly more harmful than a thorough debriefing ($\bar{X} = 3.91$), although no debriefing at all ($\bar{X} = 3.70$) did not differ from either the short or thorough debriefing. This interesting pattern in the debriefing results is evident in several measures and will be discussed later in this paper.

**Scientific Importance**

In order to insure that the scientific importance variable was correctly perceived, three separate but conceptually related measures were obtained: worth, value, and usefulness. In the first of these measures, subjects in the high scientific importance condition reported that the experiment was more worthwhile than those in the low scientific importance condition: $F (1,216) = 34.28$, $p<.001$. The other independent variables also contributed to variations in ratings of worth.

A main effect of stress was observed: $F (3,216) = 4.23$, $p<.01$. A Tukey analysis comparing the four experiments indicated that the Ax-type experiment differed reliably from the Aronson and Linder-type experiment but no differences were observed for the other two experiments. This may have been the result of
personal evaluations of the specific goals of each experiment.

Experimenter concern also produced a main effect: $F(2,216) = 5.94, p<.01$. A Tukey analysis showed subjects rated the experiment as significantly more worthwhile when a thorough debriefing was performed than when either a short or no debriefing was performed. Subjects may have interpreted the question as meaning worthwhile to themselves as well as to society. A thorough debriefing, then, would definitely be seen as more worthwhile. Subjects expect to learn as much as they can about what they have just experienced.

An interaction between stress and experimenter concern was also evident: $F(6,216) = 2.538, p<.05$. However, prior knowledge of the Milgram experiment may have confounded these results. When a post-hoc analysis was performed on those subjects who stated they had never heard of the Milgram experiment ($N = 208$), this interaction was not present.

The second check on the scientific importance variable required subjects to rate the experiment on a seven point scale ranging from "not at all valuable to science/mankind" (1) to "very valuable to science/mankind" (7). Subjects in the high scientific importance condition rated the experiments significantly higher in
value than those in the low scientific importance condition: $F(1,216) = 42.32, p<.001$.

Here again, however, the other independent variables produce main effects. The effect of stress, while present in the initial analysis with all subjects, was not evident in the analysis of those who had never heard of the Milgram experiment. Prior knowledge of an experiment seems to have differential effects on the perception of that experiment's value.

Experimenter concern also produced a main effect: $F(2,216) = 4.65, p<.05$. A Tukey test indicated that a short debriefing ($\bar{X} = 3.25$) was significantly less valuable than a thorough debriefing ($\bar{X} = 4.30$) although no debriefing at all ($\bar{X} = 3.74$) did not differ from either the short or thorough debriefing.

The third manipulation check for scientific importance was the seven point scale describing the experiment as "useful" (1) or "useless" (7). Subjects in the high scientific importance condition rated the experiment significantly more useful than subjects in the low scientific importance condition: $F(1,216) = 16.75$, $p<.001$. No effects of stress or experimenter concern were present.

Considering the large main effects of scientific importance for each of these three measures,
we may conclude that subjects did correctly perceive the manipulation.

**Experimenter Concern**

Subjects responded to the question "How concerned was the experimenter for the subject's well-being?" Ratings of concern produced a strong main effect for the extent of debriefing: $F(2,216) = 41.43$, $p<.001$. A Tukey analysis comparing the three levels indicated that no debriefing ($\bar{X} = 2.29$) did not differ significantly from a short debriefing ($\bar{X} = 2.52$) although both were reliably different from a thorough debriefing ($\bar{X} = 4.66$). This has implications for the conduct of the debriefing session in current research. It seems that only a thorough debriefing is effective in alleviating the subject's discomfort or bad feelings. The short or "standard" debriefing may only serve the purpose of allowing the experimenter to claim that he has performed one. The effect on the subject is negligible or even negative in some cases.

A main effect of scientific importance was also present in this measure: $F(1,216) = 10.67$, $p<.001$. It is possible that subjects personalized a concern for science/mankind to a concern for the subject as well.
Dependent Measures

Justifiability of Deception

It was hypothesized that subjects would rate a deception as more justifiable when the outcome of the experiment contributed to the understanding and resolution of some social problem than when the outcome had no obvious social applicability. Response to the question "How justifiable was the deception in terms of cost to you as the subject?" revealed a near significant main effect for scientific importance: $F(1,216) = 3.62$, $p<.055$. The deception was viewed as more justifiable when there was high scientific importance ($\bar{X} = 3.59$) than when there was low scientific importance ($\bar{X} = 3.16$). The potential gains from an experiment do appear to influence one's perception of the justifiability of the methods and procedures employed. The principle that the end justifies the means is evidently a widely held belief. The fact that subjects in the low scientific importance condition rated identical procedures as lower in justifiability lends support to the notion that subjects' "endorsements" of possibly unethical procedures may partially result from a benefit-oriented cover story.

Responses to this question also produced a main effect for experimenter concern: $F(2,216) = 3.69$, $p<.05$, supporting the second hypothesis of this paper.
A Tukey analysis comparing the three levels of debriefing revealed no difference between the absence of a debriefing ($\bar{X} = 3.18$) and a short debriefing ($\bar{X} = 3.14$) although both were significantly different from a thorough debriefing ($\bar{X} = 3.81$). Here again it is evident that a subject's endorsement of the methods used may be determined, in part, by the quality of the debriefing he receives.

The hypothesis that the effects of scientific importance and experimenter concern would have an effect only when the deception involved placing the subject under some form of stress was not supported. A main effect of stress was obtained with stressful experiments seen as less justifiable than non-stressful experiments: $F (3,216) = 3.60, p<.05$, but there was no evidence of interactions with the other variables. Main effects of scientific importance and experimenter concern indicate that these factors exert their influence on reports of justifiability regardless of the amount of stress involved.

A second question to measure the justifiability of the deception asked "How justifiable was the deception in terms of possible benefit to mankind?" Since the scientific importance variable was specifically designed to manipulate perception of benefit to mankind
it is only logical to expect a strong effect of scientific importance: $F (1,216) = 30.52, p < .001$.

Not as predictable, though, is an interaction between scientific importance and experimenter concern: $F (2,216) = 3.21, p < .05$. An analysis of the simple main effects revealed the source of the interaction as the effect of debriefing in the low scientific importance condition. Table 1 shows that the extent of debriefing has little effect under high scientific importance.

**TABLE 1**

Cell means for Scientific Importance X Debriefing in response to "How justifiable was the deception in terms of possible benefit to mankind?"

<table>
<thead>
<tr>
<th>DEBRIEFING</th>
<th>SCIENTIFIC IMPORTANCE</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>4.52</td>
<td>2.70</td>
<td></td>
</tr>
<tr>
<td>Short</td>
<td>3.92</td>
<td>3.22</td>
<td></td>
</tr>
<tr>
<td>Thorough</td>
<td>4.45</td>
<td>3.65</td>
<td></td>
</tr>
</tbody>
</table>

**NOTE:** The larger the number, the more justifiable the deception.

In the low scientific importance condition, the more extensive the debriefing, the greater the justifiability
of deception. One possible explanation for this may lie in the subjects' generalization of interpretation. Under conditions of low scientific importance a more extensive debriefing showing greater concern for the subject may be generalized to showing greater concern for mankind. Under conditions of high scientific importance the concern for mankind is already apparent in the description of the importance of the experiment, thus debriefing does not have as great an impact.

Perception of Deception in Terms of Ethics

Responses to the question "Is the deception unethical?" (yes or no) generated a main effect of stress: \( \chi^2(3) = 12.78, p<.01 \). A comparison of stressful experiments with non-stressful experiments revealed that the former were rated more unethical than the latter: \( \chi^2(1) = 7.97, p<.005 \).

Additionally a significant effect of scientific importance was obtained: \( \chi^2(1) = 5.43, p<.02 \). Subjects rated as more unethical those experiments suggesting low scientific importance. This is further support for the "end justifies the means" concept.

No effect of experimenter concern was present. As the question related directly to the deceptive procedures, subjects were not strongly influenced by the description of the debriefing.
Subjects Evaluation of Experiment

On a seven point scale ranging from "interesting" (1) to "dull" (7), the stress variable produced a main effect: $F (3,216) = 2.81, p<.05$. Although both stressful experiments were rated as more interesting than non-stressful experiments, a Tukey test revealed that the effect was caused by the Ax-type experiment differing significantly from the Freedman and Doob-type experiment. The other experiments did not differ markedly from either of these two. As with the "worthwhile" variable, responses concerning how interesting the experiment was may have been greatly affected by personal evaluations of the content of each experiment.

The scientific importance of the experiment also had an effect on subject's interest: $F (1,216) = 4.32, p<.05$. Subject in the high scientific importance condition found the experiments more interesting than subjects in the low scientific importance condition. It is reasonable to assume that most people view something that is worthwhile and useful as more interesting than something that has no apparent value. A valueless venture is often considered a waste of time and therefore does not generate much interest.

Willingness to Participate

The major determinant of whether or not a
subject would be willing to participate in an experiment seems to be the amount of stress involved: $F(3,216) = 20.80, p<.001$. A Tukey test comparing the four experiments indicated that subjects are more willing to participate in non-stressful experiments than stressful experiments, a finding that is not particularly surprising. Nor is it difficult to appreciate the fact that subjects are more willing to participate in scientifically important studies than those low in scientific importance: $F(1,216) = 9.93, p<.001$. An endeavor that is beneficial to society is certainly more worthy of participation and involvement than one that is not.

Subjects' willingness to participate was also a function of the extent of debriefing: $F(2,216) = 2.97, p<.06$. This marginally significant finding reaffirms the notion that subjects are concerned with the treatment they receive at the hands of the experimenter. Willingness scores for the three levels of debriefing increased as the extent of debriefing increased but a Tukey analysis indicated that only the no debriefing condition differed reliably from the thorough debriefing condition.

Importance of Experimentation in Terms of Usefulness of Results

Response to the statement "The most important aspect of an experiment involving human subjects is how
useful are the results," revealed no main effects. Subjects rated their agreement on a seven point scale ranging from "strongly agree" (1) to "strongly disagree" (7). The grand mean ($\bar{X} = 3.35$) falls between "agree somewhat" and "undecided" suggesting that usefulness of results is not an overriding concern of subjects.

A stress X debriefing interaction was found to be significant: $F(6,216) = 2.45$, $p<.05$. As evident in Table 2, the effect of debriefing is inexplicably different for each experiment. Reasons for this most probably lie in the subjects' personal evaluations of the specific content of each experiment.

**TABLE 2**

Cell means for Stress X Debriefing in response to the statement "The most important aspect of an experiment is how useful are the results."

<table>
<thead>
<tr>
<th>DEBRIEFING</th>
<th>STRESS (type of experiment)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freedman &amp; Doob</td>
</tr>
<tr>
<td></td>
<td>Aronson &amp; Linder</td>
</tr>
<tr>
<td>None</td>
<td>3.85</td>
</tr>
<tr>
<td></td>
<td>3.85</td>
</tr>
<tr>
<td></td>
<td>3.25</td>
</tr>
<tr>
<td></td>
<td>2.90</td>
</tr>
<tr>
<td>Short</td>
<td>2.85</td>
</tr>
<tr>
<td></td>
<td>3.75</td>
</tr>
<tr>
<td></td>
<td>3.20</td>
</tr>
<tr>
<td></td>
<td>3.30</td>
</tr>
<tr>
<td>Thorough</td>
<td>3.35</td>
</tr>
<tr>
<td></td>
<td>3.15</td>
</tr>
<tr>
<td></td>
<td>2.85</td>
</tr>
<tr>
<td></td>
<td>4.45</td>
</tr>
</tbody>
</table>

**NOTE:** The smaller the number, the greater the agreement with the statement.
Additionally there was an interaction between scientific importance and debriefing: $F(2,216) = 5.60$, $p < .01$. Table 3 indicates that under conditions of high scientific importance, increased amounts of debriefing lead to greater agreement with the statement. When an

TABLE 3

Cell means for Scientific Importance X Debriefing in response to the statement "The most important aspect of an experiment is how useful are the results."

<table>
<thead>
<tr>
<th>SCIENTIFIC IMPORTANCE</th>
<th>DEBRIEFING</th>
<th>High</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>3.70</td>
<td>2.97</td>
<td></td>
</tr>
<tr>
<td>Short</td>
<td>3.57</td>
<td>2.97</td>
<td></td>
</tr>
<tr>
<td>Thorough</td>
<td>3.02</td>
<td>3.87</td>
<td></td>
</tr>
</tbody>
</table>

NOTE: The smaller the number, the greater the agreement with the statement.

der endeavor is important a thorough debriefing may accentuate its value. Under conditions of low scientific importance, however, increased amounts of debriefing leads to less agreement with the statement. A thorough debriefing pointing out the subject's gullibility may be embarrassing when what he has experienced appears to be for an unimportant cause. This interpretation must be
made with caution, though, due to the presence of a triple interaction. An analysis of the simple interaction effects of scientific importance and debriefing for each of the four experiments reveals that the interaction is significant only in the Freedman and Doob-type experiment. Reasons for this are unclear.

**Importance of Experimentation in Terms of the Welfare of Subjects**

On a seven point scale ranging from "strongly agree" (1) to "strongly disagree" (7), responses to the statement "The most important aspect of an experiment involving human subjects is the safety and welfare of the subjects," produced no significant effects. The grand mean ($\bar{X} = 2.01$) reveals quite predictably that most subjects agree their safety and welfare is very important regardless of how beneficial a study might be or how extensive a debriefing they may receive.
CONCLUSION

A subject's endorsement of an experiment that has caused him some anxiety and discomfort cannot always be taken as justification for continued use of objectionable procedures. It has been demonstrated that when a subject perceives some future benefit arising from his unpleasant experiences he is more willing to offer that endorsement. There are other factors contributing to this finding such as the prestige and competence of the experimenter, and the location and sponsor of the research. Intuitively, these factors are strongly correlated with one another. A highly competent researcher would also earn his prestige and procur funds from noteworthy sources. It is also likely that he would be a resident of a prestigious university to carry out beneficial research. Further refinements to isolate precisely which variables are most instrumental in drawing out a subject's endorsement of objectionable procedures is a goal for future research. At this point one may only conclude that a prestigious researcher who can convince his subjects that what he is doing is for the good of mankind, can use procedures that for others would be out of the question.

Debriefing, too, has an impact on the subject's endorsement of an experiment, but not always in a linear
fashion. In all significant findings concerning the extent of debriefing, a short debriefing was never significantly different from no debriefing. The implications of this are quite serious. The short debriefing described in this study alluded to a personal interaction between the subject and the experimenter that lasted for ten minutes. Although this seems like a short period of time it is doubted that most real debriefings last even that long. In fact it would probably be safe to say that the majority of debriefings consist of one page xeroxed copies of technical jargon that very few subjects can relate to anyway.

A thorough debriefing does have a greater influence on subjects. In the measure of justifiability of deception, thorough debriefings resulted in significantly higher ratings of justifiability than short or no debriefings. If we are able to judge the merits of an occurrence objectively, apart from surrounding factors, it is conceivable that the procedures subjects read about should not vary significantly as a function of events happening afterwards. Reports of greater justifiability then, may indicate that subjects are reluctant to express their anger to the apologetic and well-meaning experimenter, or at the very least, are willing to forgive him his transgressions.
The hypothesis that scientific importance and experimenter concern would have an effect only when there was stress involved was not substantiated. An underestimation of subjects' concern for ethical treatment was apparently made. Stress or no stress, reactions to deceptions were affected by these other variables.

Under non-stressful conditions subjects may be less concerned with their personal comforts and more concerned with the appropriateness of deceit in an academic setting. In this situation, objections may be on a philosophical level, making determinations of the propriety of deception in general. Whereas under stressful conditions subjects are more apt to empathize with the victim of the deception and his physical discomfort, thus making determinations of the propriety of specific salient procedures.

The basic question, still left unanswered, is do the ends really justify the means? This paper is not an attempt to answer that question but rather to demonstrate that an experimenter can, if he chooses, present his "ends" in such a way that subjects will endorse his methods. Endorsement may be obtained even though his methods may be objectionable to outside observers. He may strengthen that endorsement even more by conducting a very thorough debriefing, presenting an image of himself as a concerned, feeling individual.
If, in fact, subjects honestly believe the end justifies the means, care must be taken in current and future research to insure that researchers' goals are not misrepresented merely to exact an endorsement from the subject.
REFERENCES


Ax, A. F. The physiological differentiation between fear and anger in humans. Psychosomatic Medicine, 1953, 15, 433-442.


Milgram, S. Some condition of obedience and disobedience to authority. Human Relations, 1965, 18, 57-76.


APPENDIX A

COVER SHEET

As students in Psychology you are all aware of the fact that you are required to participate as subjects in a number of experiments. The Psychology Department is planning to study the effects of certain deceptive experimental procedures on subjects. This time, rather than serve as subjects yourselves, we would like you to assist us in determining which procedures are acceptable and which procedures are not.

On the pages that follow you will find short descriptions of actual experiments that have been performed in the college setting using introductory psychology students such as yourselves. After each description there will be a short questionnaire regarding the procedures used. While answering the questions put yourself in the place of the subject described in the experiment. Answer as if you had just experienced the deceptions and the feelings involved. It is important that you do this since the results of this study may have a bearing on the experimental procedures that you may find yourself experiencing for experimental credit.

Please answer all questions even if you do not believe you have enough information to make an accurate appraisal. Remember, the subjects in these experiments
knew less than you do about what was happening. There are no "right" or "wrong" answers, we are interested in your reactions to these procedures.

Do not put your names on the questionnaire.

All responses are confidential.
APPENDIX B

INTRODUCTION TO AX-TYPE EXPERIMENT WITH DR. HOWARD BRADFORD

Dr. Howard Bradford, a well-know psychologist working at Harvard University under National Science Foundation Grant Number 1056, was conducting experiments on the psychological effects of stress on affiliation. The outcome of this particular experiment, in all likelihood, would culminate eight years of intensive research and supply Dr. Bradford with a major breakthrough in answering the questions of affiliative behavior. With the knowledge gained from this research it will be possible for people of different cultures, races, and beliefs, to live together in more harmony than has ever been experienced in the past.

As always, Dr. Bradford was very efficient in the execution of his experiments and was able to make optimum use of his time and resources. He used subjects drawn from introductory psychology courses, as he did in all previous experiments, and proceeded as follows:
Howard Bradford, a senior at Goshen College in Indiana, working with funds from the Student Research Association, was conducting experiments on the psychological effects of stress on affiliation. The outcome of this particular experiment, in all likelihood, will not contribute very much to the understanding of affiliative behavior in terms of harmony between different cultures and races, but will satisfy the requirements of one of Howard's class assignments.

As always, Howard was not very efficient in the execution of his experiments and seldom made optimum use of his time or resources. He used subjects drawn from introductory psychology courses, as he did in all previous experiments, and proceeded as follows:
Dr. Howard Bradford, a well-known psychologist working at Harvard University under National Science Foundation Grant Number 1056, was conducting experiments on the psychological factors involving obedience to authority. The outcome of this particular experiment, in all likelihood, would culminate eight years of intensive research and supply Dr. Bradford with a major breakthrough in answering the question of what leads to the blind obedience exhibited by American soldiers in MyLai in Vietnam, and by German soldiers under Hitler during World War II.

As always, Dr. Bradford was very efficient in the execution of his experiments and was able to make optimum use of his time and resources. He used subjects drawn from introductory psychology courses, as he did in all previous experiments and proceeded as follows:
Howard Bradford, a senior at Goshen College in Indiana, working with funds from the Student Research Association, was conducting experiments on the psychological factors involving obedience to authority. The outcome of this particular experiment, in all likelihood, will not contribute much to answering the question of what leads to the blind obedience exhibited by American soldiers in MyLai in Vietnam, or by German soldiers under Hitler during World War II, but will satisfy the requirements of one of Howard's class assignments.

As always, Howard was not very efficient in the execution of his experiments and seldom made optimum use of his time or resources. He used subjects drawn from introductory psychology courses, as he did in all previous experiments, and proceeded as follows:
APPENDIX F

INTRODUCTION TO FREEDMAN AND DOOB-TYPE
EXPERIMENT WITH DR. HOWARD BRADFORD

Dr. Howard Bradford, a well-known psychologist working at Harvard University under National Science Foundation Grant Number 1056, was conducting experiments on the psychological effects of feelings of deviancy on personal space. The outcome of this particular experiment, in all likelihood, will culminate eight years of intensive research and supply Dr. Bradford with a major breakthrough in finding ways of reducing the discomforts and tensions of crowded cities and enhancing the quality of life for urban residents.

As always, Dr. Bradford was very efficient in the execution of his experiments and was able to make optimum use of his time and resources. He used subjects drawn from introductory psychology courses, as he did in all previous experiments, and proceeded as follows:
INTRODUCTION TO FREEDMAN AND DOOB-TYPE EXPERIMENT WITH HOWARD BRADFORD

Howard Bradford, a senior at Goshen College in Indiana, working with funds from the Student Research Association, was conducting experiments on the psychological effects of feelings of deviancy on personal space. The outcome of this particular experiment, in all likelihood, will not contribute very much to finding ways of reducing the discomforts and tensions of crowded cities or enhancing the quality of life for urban residents but will satisfy the requirements of one of Howard's class assignments.

As always, Howard was not very efficient in the execution of his experiments and seldom made optimum use of his time or resources. He used subjects drawn from introductory psychology courses, as he did in all previous experiments, and proceeded as follows:
APPENDIX H

INTRODUCTION TO ARONSON AND LINDER-TYPE EXPERIMENT WITH DR. HOWARD BRADFORD

Dr. Howard Bradford, a well-known psychologist working at Harvard University under National Science Foundation Grant Number 1056, was conducting experiments on the psychological effects of self-esteem on impression formation. The outcome of this particular experiment, in all likelihood, will culminate eight years of intensive research and supply Dr. Bradford with a major breakthrough in understanding how people evaluate others in various types of interactions. This knowledge will be useful in industry, the academic setting, military service, in fact, in any situation where peoples' lives are influenced by the written evaluations of others.

As always, Dr. Bradford was very efficient in the execution of his experiments and was able to make optimum use of his time and resources. He used subjects drawn from introductory psychology courses, as he did in all previous experiments, and proceeded as follows:
APPENDIX I

INTRODUCTION TO ARONSON AND LINDER-TYPE EXPERIMENT WITH HOWARD BRADFORD

Howard Bradford, a senior at Goshen College in Indiana, working with funds from the Student Research Association, was conducting experiments on the psychological effects of self-esteem on impression formation. The outcome of this particular experiment, in all likelihood, will not contribute very much to the understanding of how people evaluate others in various types of interactions but will satisfy the requirements of one of Howard's class assignment.

As always, Howard was not very efficient in the execution of his experiments and seldom made optimum use of his time or resources. He used subjects drawn from introductory psychology courses, as he did in all previous experiments, and proceeded as follows:
APPENDIX J

AX-TYPE EXPERIMENT

Subjects were informed that they were participating in a data gathering study on biorhythms and electrical outputs of the body. Each subject was taken to a small room filled with an impressive array of electrical equipment where there were heavy wires cluttering the floor leading into the equipment and red warning signs on the walls. The subject was strapped into a chair and electrodes were placed on his arms and head, supposedly to record the subject's electrical output. In addition, a meter was visible to the subject with a scale ranging from 0 to 750 volts, the last 200 volts being located in a red zone on the meter.

The subject was assured that he was in no danger because the equipment merely measured his electrical output and that it was incapable of delivering any shocks. A few seconds later while the experimenter was "testing" the equipment, the lights flickered, a burst of smoke came from one of the units, and the experimenter ran from the room shouting "O my god!" At this point a small amount of electricity was delivered to the subject's forearm and the needle on the meter jumped. The lights flickered again and the current increased. This continued for several minutes with other units...
appearing to short out and the needle on the meter steadily increasing. Finally the experimenter returned, nervously apologized to the subject and unstrapped him. As they left the equipment room the experimenter requested the subject remain a few minutes to fill out some departmental forms. He then asked the subject if he would like to wait a few minutes for the forms alone or with other subjects in a nearby room. The number of subjects who preferred to wait alone was compared to a control group who did not receive any shocks.
APPENDIX K

MIGRAM-TYPE EXPERIMENT

Subjects were told they were participating in an experiment on the effects of punishment on learning. The subject was assigned the role of teacher. The learner (actually a confederate of the experimenter) was taken into the next room and strapped in an "electric chair" apparatus. The straps were to prevent excessive movement while the learner was being shocked. Electrodes were attached to the learner's wrist and electrode paste was applied to prevent blisters and burns. The teacher was told that the shocks could be extremely painful but would cause no permanent tissue damage.

The lesson consisted of a paired-associates learning task in which the teacher read a series of word pairs to the confederate learner and then read the first word of the pair along with four terms. The learner had to indicate which of the four terms had originally been paired with the first word. The teacher was instructed to shock the learner each time he gave a wrong answer. Moreover, every time a wrong answer was given the teacher had to increase the shock by 15 volts. (The shock board consisted of switches ranging from 15 to 450 volts. This range was labeled from slight shock to extremely dangerous, severe shock.)
The confederate learner's responses were predetermined by the experimenter so that successively higher shock levels up to and including 450 volts were demanded. As the shock level continued to increase the learner displayed signs of great distress, e.g., pounding on the wall and screaming. Subjects at this point turned to the experimenter for a sign to stop the experiment but were instead instructed to continue raising the shock level. The typical subject continued to raise the voltage and shock the learner. Soon, no more sounds came from the learner. The experimenter instructed the subject teacher to regard no answer as an incorrect response and continue the shocks.

Many subjects exhibited extreme signs of tension. Some were observed to sweat, tremble, stutter, bite their lips, groan and dig their fingernails into their flesh. These were characteristic rather than exceptional responses to the experiment.
APPENDIX L
FREEDMAN AND DOOB-TYPE EXPERIMENT

Ten subjects were run at a time and were led to believe the purpose of the experiment was to study impression formation. It was explained to them that to get an overall picture of their personalities, they would have to take five personality tests. The tests were supposedly scored immediately by computer and each subject was allowed to see his own personality profile. Actually all subjects received the same profile which showed them to be "different from the average person" but was not specific as to which characteristics were different. Shortly after this the experimenter brought in a sixth personality test that he had "forgotten" the first time. This test measured the subjects' feelings of alienation. The results of this test showed that everyone considered themselves slightly above average on most traits.

Under the pretext of shortage of time subjects were not able to see the results of this sixth test and were instead asked to take a seat in a large room containing many rows of chairs. The experimenter was looking at the actual distance these subjects sat from one another in the room. He would compare these distances to a similar group of subjects in the control condition.
APPENDIX M

ARONSON AND LINDER-TYPE EXPERIMENT

Subjects were told they were participating in a study of problem solving techniques. They worked in groups of four on a difficult task for about fifteen minutes. Two of the group members were actually confederates of the experimenter. One confederate left the room during the break while the other confederate suggested that one of the real subjects go into the room next door to get some more paper. Upon entering the room the subject overheard the first confederate talking with the experimenter behind a partition on the other side of the room. In that staged conversation the confederate spoke of the real subject as an intelligent and interesting person who probably had many friends.

When the subject returned to the problem solving room a second experimenter was administering a questionnaire to the others in the room. The subject was also given a questionnaire and was required to rate the others in the group on several personality characteristics. It was explained that compatibility was an important factor in speedy and accurate problem solving. The measure of interest was the difference in ratings between the subject who had overheard the conversation and the subject who had not.
APPENDIX N

QUESTIONNAIRE

1. How willing would you be to participate in an experiment of this type?

<table>
<thead>
<tr>
<th></th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>+1</th>
<th>+2</th>
<th>+3</th>
</tr>
</thead>
<tbody>
<tr>
<td>not at all willing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>very willing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. Please rate this experiment on the following bipolar scales:

<table>
<thead>
<tr>
<th></th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>+1</th>
<th>+2</th>
<th>+3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all worthwhile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>very worthwhile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>not at all valuable to science/mankind</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>very valuable to science/mankind</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>harmful</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>harmless</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>interesting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dull</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>useful</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>useless</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. Was the deception unethical in your opinion?

<table>
<thead>
<tr>
<th>YES</th>
<th>NO</th>
</tr>
</thead>
</table>

4. How justifiable was the deception in terms of cost to you as the subject?

<table>
<thead>
<tr>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>+1</th>
<th>+2</th>
<th>+3</th>
</tr>
</thead>
<tbody>
<tr>
<td>not at all</td>
<td>very justifiable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. How justifiable was the deception in terms of its possible benefit to mankind?

<table>
<thead>
<tr>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>+1</th>
<th>+2</th>
<th>+3</th>
</tr>
</thead>
<tbody>
<tr>
<td>not at all</td>
<td>very justifiable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. How concerned was the experimenter for the subject's well-being?

<table>
<thead>
<tr>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>+1</th>
<th>+2</th>
<th>+3</th>
</tr>
</thead>
<tbody>
<tr>
<td>not at all</td>
<td>very concerned</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7. The most important aspect of an experiment involving human subjects is how useful are the results.

<table>
<thead>
<tr>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>+1</th>
<th>+2</th>
<th>+3</th>
</tr>
</thead>
<tbody>
<tr>
<td>strongly agree</td>
<td>agree</td>
<td>undecided</td>
<td>disagree</td>
<td>strongly disagree</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. The most important aspect of an experiment involving human subjects is the safety and welfare of the subjects.

<table>
<thead>
<tr>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>+1</th>
<th>+2</th>
<th>+3</th>
</tr>
</thead>
<tbody>
<tr>
<td>strongly agree</td>
<td>agree</td>
<td>undecided</td>
<td>disagree</td>
<td>strongly disagree</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>