

# Anticipating Informed Consent

## *An Empirical Approach*

ELLEN BERSCHEID *University of Minnesota*<sup>1</sup>  
ROBERT STEVEN BARON *University of Iowa*  
MARSHALL DERMER *University of Minnesota*<sup>2</sup>  
MARK LIBMAN *University of Minnesota*<sup>3</sup>

Section 3-4, "Obtaining Informed Consent to Participate," of the new Ethical Standards for Research with Human Subjects, recently adopted by the American Psychological Association (Cook et al., 1972) begins:

The psychologist's ethical obligation to use people as research participants only if they give their informed consent rests on well-established traditions of research ethics and on strong rational grounds. The individual's human right of free choice requires that his decision to participate be made in the light of adequate and accurate information [p. VI].

The principle of informed consent has long guided much psychological research. A prerequisite for obtaining funds from the U.S. Public Health Service (1969), for example, has been a prior review of proposed research by an appropriate committee of the applicant institution. Among other considerations, the review is to assure that "the methods used to obtain informed consent are adequate and appropriate [p. 1]."

In conflict with the principle of informed consent is the practice of misleading prospective research participants as to the true purpose of the experiment and withholding from them information concerning the exact procedures to be followed during the course of their participation. "Deception"

methodology, however, often serves to facilitate the manipulation of independent variables that otherwise would not be amenable to rigorous investigation and to minimize the extent to which "demand characteristics" (Orne, 1964) are operative in experimental settings.

The new APA Ethical Standards (1972) document recognizes that methodological requirements frequently present the investigator with ethical dilemmas. As a guideline for resolution in the case of deception methodology, the document outlines a "number of considerations that *may* make the use of deception more acceptable [p. VII]": The research problem is of great importance; it may be demonstrated that the research objectives cannot be realized without deception; the subject is allowed to withdraw from the study at any time and to withdraw his data if he so wishes when the concealment or misrepresentation is revealed to him; the investigator takes full responsibility for detecting and removing stressful aftereffects; and finally,

there is sufficient reason for the concealment or misrepresentation that when the research participant is later informed . . . , he can be expected to find it reasonable, and so suffer no loss of confidence in the integrity of the investigator or others involved . . . [p. VII].

The USPHS (1969) directive also appears to make allowance for research in which "informed consent . . . can be anticipated [p. 10]."

How can one determine in advance of experimentation whether a potential subject would consent to participation if he were informed of the aims and procedures involved in the experimentation? Under what circumstances is he likely to later find the concealment and deception reasonable? A

---

<sup>1</sup> This research was funded in part by National Institute of Mental Health Grant MH 16729 to Ellen Berscheid. The authors wish to express their appreciation of Kinley Larantz' statistical advice.

Requests for reprints should be sent to Ellen Berscheid, Department of Psychology, University of Minnesota, Minneapolis, Minnesota 55455.

<sup>2</sup> Now at University of Wisconsin, Milwaukee.

<sup>3</sup> Now at York University, Toronto, Ontario, Canada.

number of strategies may be followed to answer these questions.

Perhaps the most usual procedure for determining whether consent may be anticipated is for the investigator himself to estimate whether his potential subjects would consent to the experimentation if given full procedural information. The potential pitfalls of this strategy need not be outlined.

Another current practice is to present the proposed experimental procedure, along with a description of the subject population, to an "ethics advisory group" in order to obtain their collective opinion as to whether participants are likely to find the concealment reasonable. While such judgments may not be as vulnerable to self-interest as those of the principal investigator, a number of questions about the predictive accuracy of any given "advisory" group could be raised.

We propose that a simple, but more accurate, alternative to the strategies outlined above might be to draw a sample from the proposed subject population, present it with the full procedure to be followed in the experiment along with the purpose of the experimentation, and determine the extent to which these subjects would be willing to participate in the experiment described. Baumrind (1971, p. 889) has recently suggested a similar strategy. From this "role-play-sampling" procedure, consent rates could be projected for the subject population. In addition, the degree to which alternative procedures (perhaps equally desirable on methodological grounds) would be acceptable to potential subjects could be easily and inexpensively determined prior to actual participation in experimentation by any human subject.

The proposed strategy is admittedly a "role-playing" procedure. As such, it may be liable to several of the criticisms that have been lodged against role-play methodology with regard to its suitability as a substitute to the nonvicarious experiencing of an experimental situation (Freedman, 1969). It has been charged, for example, that when a person is asked to imagine himself in a situation, his behavior may diverge considerably from his behavior when actually placed in that situation. With respect to the role-play-sampling procedure, it is clear that a prospective participant may not always accurately foretell the degree of discomfort or stress that will be aroused by the experiment even when it is carefully described in detail. It is possible, in addition, that he may

entertain self-enhancing projections of the behavior he is likely to exhibit in the situation. Unfortunately, these problems are also inherent in the principle of informed consent. Adoption of this principle alone does not guarantee that an individual who has given fully informed consent to participate in an experimental procedure will not later regret his decision due to the inaccuracy of his subjective judgments made prior to consent. For this reason it should be emphasized that the proposed procedure is addressed only to the question of consent which plays a central, but by no means exclusive, role in the attempt to protect the subject's welfare.

In the present study we wished to investigate the degree to which "informed consent" to a number of published social-psychological experiments could have been anticipated had a sampling procedure for determining likelihood of consent been used prior to experimentation. To meet the criticism that role-play subjects, unlike potential participants, may fail to become "involved" since their decisions are without consequence, two variants of the role-play consent procedure were investigated. A high-consequence condition was constructed in an attempt to approximate an actual consent situation in which subjects believe their consent responses will determine actual experimental participation. In a low-consequence condition, responses were obtained under the more usual role-playing circumstance in which subjects anticipate that their decisions are largely academic and will most probably never have any real consequence for them.

The subject sample was drawn from the population most typically used in social-psychological experimentation, the lower-division college student. Included for investigation of prospective subjects' reactions were both those experiments that might be considered innocuous in terms of subject discomfort (nonstressful) and those that we suspected to have been stressful to subjects during the course of experimentation (stressful). It was hypothesized that subjects in the high-consequence condition would respond more cautiously to stressful than to nonstressful experiments, while subjects in the low-consequence condition would not respond as differently to the two types of experiments.

Use of the proposed strategy for determining whether consent may be anticipated necessitates a judgment of whether the consent that follows the receipt of a particular body of information should

be considered "informed" or "uninformed." What are the relevant pieces of information that should be made available to prospective participants in order for their consent to be deemed informed? Principle 3 of the APA Standards (Cook et al., 1972) document states, in part, "The investigator should inform the participant of all features of the research that reasonably might be expected to influence willingness to participate [p. VI]." The influence of various types of information upon willingness to consent was investigated in the present study. Stressful and nonstressful experiments were described to potential subjects at one of seven levels of information:

(a) *Rationale* (cover story information only). A glance at any recent journal documents the fact that potential volunteers tend to consent to participation in studies even when they are given only a cursory rationale for their participation. It was expected that this high willingness to participate in experimentation would be reproduced when only the "cover stories" which were used in each of the experiments were presented to the subjects.

(b) *Procedure* (procedural information presented after cover story). In the case of stressful experiments, it was expected that additional knowledge of procedure should cause a significant drop in level of consent. A significant drop was also predicted for the nonstressful experiments, although the decrease was expected to be less than in the case of stressful experiments. In making the latter prediction it was reasoned that the effect of rendering consent on the basis of limited information is probably due not only to an implicit trust in the experimenter, but also to the surrender of responsibility for the decision as a result of the unequal power inherent in the prospective volunteer's lack of information. If more information is provided, even if that additional procedural information no more than reinforces the initial rationale, the potential volunteer might take more responsibility for his decision to consent, and, hence, consent level should be lowered even in the case of nonstressful experiments.

(c) *Desirable behavioral expectation* versus (d) *undesirable behavioral expectation* (expectation information presented after cover story and procedure). The potential subject's subjective prediction of whether he would be likely to behave in a "normal," justifiable, or socially desirable manner during the course of the experiment might affect

likelihood of consent. Most usually, of course, this information is not accessible to the investigator to give to subjects prior to experimentation since the question of how subjects will behave is usually the very purpose of the experimentation. Nevertheless, we thought it interesting to investigate the impact of this variable on consent, since it appears that it is partially on these grounds that the ethicality of several experiments has been called into question. For example, one can question the extent to which the Milgram (1963) experiment would have received the adverse publicity and criticism it received if most of his subjects had defied the experimenter early in the procedure. Within the present study, rather than leave it to the individual subject's imagination as to how he would behave in each experimental context, his expectation that he would be likely to engage in either socially desirable or undesirable behavior was directly manipulated.

(e) *Neutral debrief* (debriefing information presented after cover story, and procedure), (f) *desirable debrief* (debriefing information presented after cover story, procedure, and desirable behavioral expectation information), and (g) *undesirable debrief* (debriefing information presented after cover story, procedure, and undesirable behavioral expectation information). Since some of the information that is typically contained within a debriefing session can be hypothesized to affect a prospective volunteer's decision to consent, three debriefing conditions were included for low-consequence subjects. Debriefing information was withheld from high-consequence subjects because it was feared that such information might jeopardize the credibility of the high-consequence manipulation.

Debriefing sessions usually provide the participant with some insight into the problem which has been studied and the purposes of the research in which he has participated. In experiments employing deception methodology, a participant also learns that he was deliberately misled as to the true purpose of the experiment and that his prior consent to participate was not fully informed. While it was hypothesized here that the receipt of this information would increase willingness to consent, the effects of debriefing have not always been assumed to be ameliorative (e.g., Baumrind, 1964).

To test the hypotheses outlined above and to explore the feasibility of the proposed "role-play-

sampling" procedure, subjects were randomly assigned to 1 of 11 conditions. Eight conditions were defined by crossing consequence level (high versus low) with nature of information (rationale, procedure, desirable behavioral expectation, undesirable behavioral expectation). In addition, the three debrief conditions outlined above were run at the low-consequence level. Crossed with the above between-subject factors was a within-subject stress factor such that subjects read descriptions of both stressful and nonstressful experiments within each of the 11 conditions.

## Method

### SUBJECTS

The subjects were 106 undergraduates, 38 males and 68 females, enrolled in an introductory psychology course at the University of Minnesota, who participated for course credit. In order to qualify for participation, it was necessary that each subject commit himself to attend a "second session" the following week for which he would receive \$2. Commitment to a second session was necessary to establish the credibility of the high-consequence manipulation.

### PROCEDURE

The subjects arrived at the assigned room at various times during the day. There were typically 5-10 subjects working on test booklets at long desks at any one time. The subjects were randomly assigned to booklets; the contents of the booklet determined the experimental condition. No verbal instructions were given by the experimenter. The subjects were not aware of their participation in a multi-dimensional study since a filler task was included at the end of the booklet so that the task required the same amount of time to complete in all conditions. When the subject had worked for approximately 50 minutes, the experimenter interrupted him and asked him to self-address an envelope so that he could be contacted for the second session. The subject was then free to leave. The envelope was later used to mail a description of the aims of the present study to the subjects and to relieve them of the necessity of a second session.

The first page of each booklet contained an introduction to the study and instructions. All of the subjects were informed that the study had two purposes, one of which was "to investigate how people feel about participating in various types of psychological experiments."

### CONSEQUENCE MANIPULATION

*High-consequence-condition* subjects were told that the second purpose of the study was "to allow you to exercise some choice as to the experiment in which you will be a subject." In addition, these subjects were told, "We would like to know how *you* would *personally* feel if you had participated as a subject in each experiment. We will use some of your answers (e.g., the question 'How willing would you be . . . ?') to determine in which of these experiments you will later participate as a subject."

*Low-consequence-condition* subjects were told that the second purpose of the study was to use their responses to the current booklet to help interpret additional "judgments you will make on a series of experiments next week."

### BACKGROUND INFORMATION

Before reading the experimental descriptions, the subjects were asked to answer a series of items that provided background information. The sex and major of the respondent, as well as the number of previous psychological studies in which he had participated, were assessed. In addition, 16 items from the Berger (1952) Acceptance of Self Scale were administered. The items were chosen from the "self" subscale that presumably measures chronic self-esteem.

After responding to these items, the subjects moved on to a series of bipolar adjectives designed to measure their attitudes toward psychology and psychologists (e.g., trustworthy-untrustworthy, rewarding-punishing) and experimentation (e.g., interesting-dull, pleasant-unpleasant).

### ROLE-PLAYING INSTRUCTIONS AND RESPONSES

After completing this section, the subjects read six experimental descriptions which were presented in random order. The importance of projection into these experimental situations was stressed. To aid the subjects in this effort, all were instructed that:

Details will be given in the descriptions to help you to imagine yourself in the settings. For example, reference might be made to a typical subject—in this case you are to imagine that *you are that typical subject*, and that his behaviors and feelings are your own. After reading about an experiment please answer the questions about it in the separate answer booklet. Remember to answer these questions *as if you had just participated in the experiment*.

After reading each description, the subjects answered a series of items designed to measure their reactions to the experiment. Among these was the response measure of major interest: "How willing would you have been to participate in this experiment?" Included also were bipolar adjective scales designed to assess the subject's estimate of his well-being following the experiment and the favorability of his evaluation of the experiment itself.

### EXPERIMENTAL DESCRIPTIONS

Choice of the six experiments included in the present study was arbitrary, but the following were thought to be fairly representative of a range of experiments that have appeared in the social-psychological literature in the past 10 years. The first four were considered to be stressful, each for different reasons; the last two were thought to be relatively nonstressful.

*Milgram (1963)*. This study of obedience in a harm-doing situation was included since it has probably received more public criticism on ethical grounds than any other piece of work in social psychology. The majority of Milgram's subjects continued to pull a lever which they believed administered electric shock to a victim to the point at which they were ostensibly delivering dangerous levels of punishment.

*Berscheid, Boye, and Walster (1968)*. This study used harm-doing as an independent rather than a dependent variable. The subjects tended to derogate "victims" whom

they had ostensibly harmed through administration of electric shock and from whom they expected no retaliation.

*Barber and Hahn (1962)*. The control condition from this study of hypnotic pain reduction was used since it involved actual physical pain. The subjects were required to immerse one of their hands in painfully cold water for three minutes.

*Orne and Scheibe (1964)*. This procedure was adopted from Orne and Scheibe's "demand" condition in which the subjects responded to danger cues by interpreting a harmless situation as a very threatening one.

*Festinger and Maccoby (1964)*. This was a study of the effect of distraction on attitude change. A persuasive communication was presented while the subjects watched an amusing film.

*Deutsch and Krauss (1960)*. The subjects were allowed to compete or cooperate in a simulated bargaining situation involving the movement of toy trucks.

#### INFORMATION MANIPULATION

As previously noted, seven information conditions were employed. To illustrate each, we present below the complete text used to describe the Milgram (1963) experiment:

(a) *Rationale*. Only a typical cover story was given:

This experiment was designed to study the effects of punishment on learning. We know *very little* about this topic because almost no truly scientific studies have been made of it in human beings. For instance, we don't know how *much* punishment is best for learning—and we don't know how much difference it makes as to who is giving the punishment, whether an adult learns best from a younger or an older person than himself—or many things of that sort.

So in this study we are bringing together a number of students and asking some of them to be teachers and some of them to be learners. We want to find out just what effect different people have on each other as teachers and learners, and also what effect *punishment* will have on learning in this situation.

(b) *Procedure*. In this condition, the subjects received both the rationale and information regarding the experimental apparatus and procedure. That is, in addition to the information above, the subjects read the following:

You were randomly assigned the role of teacher. You met another student who was randomly assigned the role of learner. The learner 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. It would be impossible for him to escape the shock. An electrode was attached to the learner's wrist and electrode paste was applied to prevent blisters and burns. Although the shock could be extremely painful, it could cause no permanent tissue damage.

The lesson consisted of a paired-associates learning task. You read a series of word pairs to the learner and then read the first word of the pair along with four terms. The learner indicated which of the four terms had originally been paired with the first word. He communicated his answer by pressing one of four switches in front of him which in turn lit one of four numbered squares in an answer box on a shock generator in front of you.

The shock generator had an on-off switch and 30 lever switches to set the voltage level of the shock. Voltage

ranged from 15 to 450 volts. This resulted in a range from slight shock to extremely dangerous, severe shock.

When you depressed a switch a pilot light corresponding to each switch was illuminated in bright red, and electric buzzing was heard, an electric blue light labeled "voltage energizer" flashed, the dial on the voltage meter swung to the right and various relay clicks sounded.

You were told to administer a shock to the learner each time he gave a wrong response. Moreover, you were told to move the voltage one level higher on the shock generator each time the learner gave a wrong answer. Before administering the shock you announced the voltage level. The voltage level started at 15 volts and you were told to increase it one step (i.e., 15 volts) each time the learner gave a wrong answer. The experimental session would end when the learner had mastered 10 pairs. Before beginning you were given a sample shock at the third voltage level, i.e., 45 volts, to give you an idea of the intensity of the shock. The learning trials were then begun. Typically your learner had great difficulty mastering the task and successively higher shock levels up to and including 450 volts were demanded.

(c) *Desirable behavioral expectation*. The subjects in this condition read all of the information above and, in addition, learned that:

You as the typical teacher initially raised the voltage and shocked the learner for each mistake. The learner gave no subjective reaction to the shock up to 285 volts. When the shock reached around 300 volts, however, the learner displayed signs of great distress, e.g., pounding on the wall of the room in which he was bound to the electric chair. You were able to hear these reactions. At this point, as a typical subject, you refused to continue in spite of the experimenter's urgings that cooperation was essential.

(d) *Undesirable behavioral expectation*. The subjects in this condition received information identical to that of Condition b and, in addition, read the following:

As the typical teacher, you continued to raise the voltage and shock the learner. The learner gave no reaction to the shock up to 285 volts. When the shock reached about 300 volts, however, the learner displayed signs of great distress, e.g., pounding on the wall of the room in which he was bound to the electric chair. You were able to hear these reactions. After this outburst no more responses were given by the learner but this absence of responding in itself constituted an error. You continued to raise the voltage level. At around 315 volts the learner was again heard in great agitation but after this the learner fell silent. But you as the teacher continued to increase the voltage in steps until 450 volts were being given. This level was labeled as "extremely dangerous severe shock." You displayed very noticeable distress as you continued to shock the learner. You twitched, trembled, and laughed uncontrollably in a bizarre fashion. You were aware that you were giving extremely painful shocks. You appealed to the experimenter to terminate the procedure but he told you to continue since your cooperation was absolutely essential. You, as the typical teacher, continued to increase the voltage and shock the learner until the end of the scale—450 volts—was reached.

(e) *Neutral debrief*. The subjects in this condition received the information contained in Condition b and, in addition, the following debriefing information:

After the experiment you were informed of the true purpose of the study. The experimenter was really measuring how long and to what voltage level you would continue to shock a victim. This was part of a larger study investigating some of the conditions under which people would obey orders. Many other subjects had participated and had responded in the same manner as you did. The study was part of a long-overdue investigation of the common belief that Americans are *not* blindly obedient. No shock was actually being given since the shock generator was not really connected to the electrodes on the learner. The learner was revealed to be an employee and confederate of the experimenter. You were introduced to your "victim" and assured that you had caused no harm whatsoever. Every effort was made to assure you that you left the situation with a complete insight into the purpose and nature of the experiment and that you left in a state of well-being.

This condition differed from Conditions *f* and *g* below in that the subject was given no information concerning how he would have been likely to behave during the course of the experiment and presumably reached his own conclusions.

(*f*) *Desirable debrief*. These subjects received the information provided in Condition *c* and, in addition, the same debriefing information provided in Condition *e*.

(*g*) *Undesirable debrief*. The subjects received the information provided in Condition *d* and, in addition, the same debriefing information provided in Condition *e*.

## Results and Discussion

### WILLINGNESS TO PARTICIPATE

The primary dependent measure, willingness to participate in a particular experiment, was operationalized as the subject's response to the following question: "Assuming that you would have reacted as described and that it had been possible for you to know everything about this experiment that you now know, would you have chosen to participate in it?" The response categories, "definitely not," "probably not," "maybe," "probably," and "definitely," were respectively scored from 1 through 5. Responses were averaged for the two nonstressful and also for the four stressful experiments.

It will be recalled that it was hypothesized that high-consequence as compared to low-consequence subjects would respond more cautiously to the experiments since the former were instructed so that their responses would help determine in which study they would subsequently participate. In order to assess the extent to which the consequence and information (between-subject) factors, together with the stress (within-subject) factor, affected willingness to participate, a 2 (between)  $\times$  1 (within) analysis of variance was performed on the "willingness to participate" data. This analysis revealed neither a consequence main effect ( $F =$

.004,  $df = 1/68$ ,  $p = .95$ ), nor did it indicate that consequence interacted with the other variables (Consequence  $\times$  Information,  $F = .53$ ,  $df = 3/68$ ,  $p = .66$ ; Consequence  $\times$  Stress,  $F = .056$ ,  $df = 1/68$ ,  $p = .81$ ; Consequence  $\times$  Stress  $\times$  Information,  $F = .64$ ,  $df = 3/68$ ,  $p = .59$ ).

The inclusion of the consequence factor was an attempt to approximate an actual consent situation in which the prospective subject believes his response will determine participation. In the absence of a consequence main effect and related interaction effects, it is especially unfortunate that we have no means of assessing the absolute level of consequence, or "involvement," that our data reflect. It must be assumed that this level is somewhat less than would have been the case had the subjects been placed in an actual consent situation. Despite the necessity of adopting a conservative view, evidence, which will be presented shortly, indicates that the subjects did take the tasks seriously; their responses do not at all appear to be thoughtless, indiscriminating, or particularly accepting. In any event, the absence of consequence main effects and interactions, on all dependent measures examined, was capitalized on by collapsing across this factor in the analyses which follow.

It was predicted that both information, a between-subject factor, and the stress associated with the experiments, a within-subject factor, would affect consent. The relevant means for each condition are presented in Table 1. An unweighted means analysis of variance performed on these data yielded both significant information ( $F = 3.08$ ,  $df = 6/99$ ,  $p = .008$ ) and stress ( $F = 90.40$ ,  $df = 1/99$ ,  $p < .00001$ ) main effects. In addition, the interaction between these two factors was significant ( $F = 3.22$ ,  $df = 6/99$ ,  $p = .006$ ).

Tests for the simple effects of stress within information conditions revealed that subjects were reliably more willing to participate in nonstressful as compared to stressful experiments with the exception of the rationale ( $t = .57$ ,  $df = 99$ ,  $p = .57$ ) and desirable debrief ( $t = 1.60$ ,  $df = 99$ ,  $p = .11$ ) conditions. The absence of a stress effect within the rationale condition is explicable in terms of the paucity of stress information within this condition. The experiments were, however, described extensively in the desirable behavioral expectation and desirable debrief conditions. The presence of a stress effect in the former condition

TABLE 1

*"Would You Have Chosen to Participate" Means*

Type of experiment	Information level						
	Rationale (n = 19)	Procedure (n = 18)	Desirable behavioral expectation (n = 17)	Undesirable behavioral expectation (n = 22)	Neutral debrief (n = 8)	Desirable debrief (n = 11)	Undesirable debrief (n = 11)
Nonstressful <sup>a</sup>	3.79 <sub>d</sub>	3.67 <sub>d</sub>	4.24 <sub>d</sub>	3.91 <sub>d</sub>	4.19 <sub>d</sub>	3.59 <sub>d</sub>	4.05 <sub>d</sub>
Stressful <sup>b</sup>	3.57 <sub>od</sub>	2.42 <sub>ab</sub>	2.80 <sub>abe</sub>	2.29 <sub>a</sub>	3.41 <sub>bc</sub>	2.97 <sub>abcd</sub>	3.10 <sub>abc</sub>

Note. 1 ≈ "definitely not," 2 ≈ "probably not," 3 ≈ "maybe," 4 ≈ "probably," 5 ≈ "definitely." Means with common subscripts do not differ, at the .05 level, as determined by two-tailed *t* tests between stress conditions and Duncan multiple-range tests within stress conditions.

<sup>a</sup> Each cell mean in this row is the average (across subjects) of each subject's average response to two nonstressful experiments.

<sup>b</sup> Each cell mean in this row is the average (across subjects) of each subject's average response to four stressful experiments.

and absence of such an effect in the latter condition suggest that debriefing information did exert a slight ameliorative effect. The moderating effect of debriefing information was restricted, however, to the desirable debrief condition (see Table 1).

To examine the simple effects of information within stress conditions, Duncan multiple-range tests were performed (see Table 1). These analyses revealed no significant differences due to information within the nonstressful conditions. Within the stressful conditions, however, type of information did affect willingness to participate. Willingness to participate was significantly higher in the rationale condition than it was in either the procedure or undesirable behavioral expectation conditions. These differences again reflect the paucity of stress information within the rationale condition as compared to the other conditions. Willingness to participate was also significantly higher for the subjects who were never sensitized to undesirable behaviors and who knew the purpose behind each experiment (neutral debrief), as compared to the subjects who were sensitized to undesirable behaviors but did not receive the debriefing information (undesirable behavioral expectation).

The Duncan tests for the simple effects of infor-

mation within the stressful condition also revealed a slight tendency for debriefing information to increase consent, such that the neutral, desirable, and undesirable debrief conditions did not differ significantly from the rationale condition. However, levels of consent within the former conditions did not differ from those found in the procedure, desirable behavioral expectation, and undesirable behavioral expectation conditions where debriefing information was absent. Thus, debriefing information was again found not especially effective in increasing consent levels for stressful experiments.

In the preceding analyses, consent was treated as a continuous variable. Within discussions of informed consent, however, it is often conceptualized as a dichotomous variable: A subject either agrees or refuses to participate. To investigate how each of the experiments would have fared if the present role-play procedure had been used prior to actual experimentation and a "percentage of refusals" decision rule had been formulated and enforced, responses to the willingness-to-participate question were rescaled such that "definitely not" responses were considered instances of consent refusal while the remaining responses were considered expressions of consent. These data are presented in Table 2,

TABLE 2

*Withholding of Consent Expressed as a Percentage of Decisions*

Type of experiment	Information level						
	Rationale	Procedure	Desirable behavioral expectation	Undesirable behavioral expectation	Neutral debrief	Desirable debrief	Undesirable debrief
Nonstressful	3	3	0	0	0	0	0
Stressful	3	35	28	42	9	7	16

Note. Observations within any one cell are dependent. Observations reflecting the simple effects of stress within a particular level of information are dependent.

TABLE 3

## "Personal Reaction following Experiment" Means

Type of experiment	Information level						
	Rationale (n = 19)	Procedure (n = 18)	Desirable behavioral expectation (n = 17)	Undesirable behavioral expectation (n = 22)	Neutral debrief (n = 8)	Desirable debrief (n = 11)	Undesirable debrief (n = 11)
Nonstressful <sup>a</sup>	11.93 <sub>bc</sub>	11.37 <sub>c</sub>	12.41 <sub>c</sub>	9.98 <sub>c</sub>	11.50 <sub>c</sub>	10.87 <sub>c</sub>	11.09 <sub>c</sub>
Stressful <sup>b</sup>	11.20 <sub>b</sub>	8.12 <sub>ab</sub>	8.75 <sub>ab</sub>	7.10 <sub>a</sub>	10.00 <sub>abc</sub>	9.86 <sub>abc</sub>	9.62 <sub>abc</sub>

Note. 2 ≈ "sadder/less satisfied than before," 10.0 = scale midpoint, 18 ≈ "happier/more satisfied than before." Means with common subscripts do not differ, at the .05 level, as determined by two-tailed *t* tests between stress conditions and Duncan multiple-range tests within stress conditions.

<sup>a</sup> Each cell mean in this row is the average (across subjects) of each subject's average response to two nonstressful experiments.

<sup>b</sup> Each cell mean in this row is the average (across subjects) of each subject's average response to four stressful experiments.

where each figure represents the percentage of consent refusal decisions for each condition. For example, in the nonstressful rationale cell there were 19 subjects who judged two experiments, yielding a total of 38 decisions. Of these 38 decisions, 1 was indicative of refusal. Thus,  $1/38 \times 100 = .03 \times 100 = 3\%$ .

Unfortunately, a conceptually meaningful analysis of this measure was difficult due to the repeated-measures design. Analyses that met statistical canons required further rescaling, with the consequence that the analyses lost substantive import. It will be noted, however, that the results presented in Table 2 are similar in pattern to those presented in Table 1,<sup>4</sup> with the exception that when the data are viewed from this perspective, the effects seem substantially more robust within the stressful condition. Thus, in the stressful cells, procedural information produced more than 10 times the refusals (35% refusals) than rationale information alone (3% refusals). Moreover, within the stressful cells a comparison of the refusal rates in the two behavioral expectation conditions to those in the procedure cell shows that when the subjects were given desirable behavioral expectations, their refusal rate dropped by one-fifth (to 28%) and that when they had undesirable expect-

tations, refusal rates increased by one-fifth (to 42%).

Finally, the effects of debriefing seem noticeably more powerful when one focuses on definite refusals. The pooled rate of refusal for the procedure, desirable behavioral expectation, and undesirable behavioral expectation conditions is 35%. The pooled rate of refusal in the three debrief conditions is 11%. This represents a two-thirds reduction in refusals as a result of debriefing information. While debriefing did not reduce refusals to the level observed in the rationale cells (3%), these data suggest that the effects of the information treatments might be more substantial than examination of the continuous "willingness" data indicates.

Baumrind (1971) has suggested that if more than 5 out of 100 pretest subjects question an experimental procedure the study should not be run. This type of rule considers only subject risk, of course, and does not take into consideration the potential merit of the experimentation. It is interesting nonetheless that application of such a rule to our anticipated informed consent rates would result in the stressful experiments not being run; when procedural information about the stressful experiments was provided, more than 5% of the subjects refused consent. This remained true even in instances in which debriefing information was additionally included. The nonstressful experiments, however, passed the 5% criterion.

## PERSONAL REACTIONS

Baumrind (1964) has speculated that debriefing information may embarrass the subject and result in a loss of self-esteem. To tap this dimension, subjects were asked after the description of each ex-

<sup>4</sup> The high-consent refusal rates within the stressful conditions, when considered together with the presence of reliable information and interaction effects in previous analyses, underscore and provide a conservative demonstration of the general caution with which subjects approached the role-playing task. We could have made subjects appear even more cautious by dichotomizing the original measure such that both "definitely not" and "probably not" responses were scaled as instances of consent refusal. This scaling problem can, of course, be avoided in future studies by providing only yes versus no as response alternatives.

periment: "If you had participated as a subject in this particular experiment, how do you think you would have felt after it was over?" Responses were collected on two 9-point scales ( $r = .75$ ) which ranged from "Happier than before" to "Sadder than before," and "More satisfied with yourself than before" to "Less satisfied with yourself than before." Responses to these two scales were summed and averaged for the two nonstressful and also for the four stressful experiments. A score of 2 corresponds to "sadder/less satisfied," while a score of 18 corresponds to "happier/more satisfied." The relevant means for each condition are presented in Table 3.

An unweighted-means analysis of variance yielded both significant information ( $F = 3.64$ ,  $df = 6/99$ ,  $p = .003$ ) and stress effects ( $F = 74.34$ ,  $df = 1/99$ ,  $p < .00001$ ), as well as a significant interaction effect ( $F = 3.39$ ,  $df = 6/99$ ,  $p = .004$ ).

Tests for the simple effects of stress indicated that subjects reliably expressed more favorable reactions to the nonstressful as compared to the stressful experiments, except in the rationale ( $t = .81$ ,  $df = 99$ ,  $p = .41$ ) and the three debrief conditions (neutral debrief:  $t = 1.67$ ,  $df = 99$ ,  $p = .10$ ; desirable debrief:  $t = 1.12$ ,  $df = 99$ ,  $p = .27$ ; undesirable debrief:  $t = 1.64$ ,  $df = 99$ ,  $p = .10$ ). The presence of a stress effect within the procedure, desirable behavioral expectation, and undesirable behavioral expectation conditions, together with the absence of such an effect when the debriefing information was included, suggests that debriefing information increased the favorability of personal reaction. The debriefing response means are at approximately the scale midpoint, which indicates that the subjects believed they would feel about the same after the experiments as they did before the experiments.

Looking at the differences between information levels within stress conditions, Duncan multiple-range tests again revealed no significant differences due to information within the nonstressful conditions (see Table 3). Within the stressful conditions, however, one significant difference due to information level emerged. As before, the subjects within the rationale condition responded more favorably than the subjects in the undesirable behavioral expectation condition.

#### EVALUATION OF EXPERIMENTS

The subjects evaluated each experiment by responding to the following question: "If you had participated as a subject in this particular experiment, how do you think you would rate it with respect to the following dimensions?" Responses were collected on two 9-point scales ( $r = .85$ ) which ranged from "worthwhile" to "not worthwhile," and "valuable to science" to "not valuable to science." These responses were combined, such that the final composite ranged from 2, "not worthwhile/not valuable," to 18, "worthwhile/valuable." The relevant means are presented in Table 4.

An analysis of variance again revealed significant information ( $F = 4.82$ ,  $df = 6/99$ ,  $p = .0002$ ), stress ( $F = 7.98$ ,  $df = 1/99$ ,  $p = .006$ ), and interaction ( $F = 3.85$ ,  $df = 1/99$ ,  $p = .002$ ) effects. The latter effect, however, was due to the presence of a stress effect within the undesirable behavioral expectation condition, and the absence of this effect within the remaining information conditions. No significant differences in evaluation between information levels within the nonstressful conditions were found (see Table 4). Within the stressful conditions, however, level of information did affect evaluation. Examination of Table 4 will reveal

TABLE 4

#### "Evaluation of Experiments" Means

Type of experiment	Information level						
	Rationale ( $n = 19$ )	Procedure ( $n = 18$ )	Desirable behavioral expectation ( $n = 17$ )	Undesirable behavioral expectation ( $n = 22$ )	Neutral debrief ( $n = 8$ )	Desirable debrief ( $n = 11$ )	Undesirable debrief ( $n = 11$ )
Nonstressful <sup>a</sup>	15.05 <sub>od</sub>	12.47 <sub>o</sub>	12.59 <sub>o</sub>	13.89 <sub>o</sub>	13.57 <sub>o</sub>	11.27 <sub>o</sub>	12.05 <sub>o</sub>
Stressful <sup>b</sup>	15.73 <sub>bd</sub>	10.01 <sub>bc</sub>	11.21 <sub>bc</sub>	9.85 <sub>b</sub>	13.60 <sub>abc</sub>	12.12 <sub>abc</sub>	12.66 <sub>abc</sub>

Notes. 2 ≈ "not worthwhile/not valuable," 10 = scale midpoint, 18 ≈ "worthwhile/valuable." Means with common subscripts do not differ, at the .05 level, as determined by two-tailed  $t$  tests between stress conditions and Duncan multiple-range tests within stress conditions.

<sup>a</sup> Each cell mean in this row is the average (across subjects) of each subject's average response to two nonstressful experiments.

<sup>b</sup> Each cell mean in this row is the average (across subjects) of each subject's average response to four stressful experiments.

that evaluations were reliably more favorable in the rationale condition than in either the procedure, desirable behavioral expectation, or undesirable behavioral expectation conditions.

To summarize the findings across dependent measures, responses to nonstressful experiments were unaffected by variations in the amount and nature of experimental information, while stressful experiments were substantially affected by this manipulation. Focusing on stressful experiments, the most favorable responses were observed within the rationale condition, while the least favorable responses occurred within the undesirable behavioral expectation condition, with the differences between these conditions being reliable on each dependent measure. The favorability of responses within the former condition suggests that students, in the absence of additional information, hold a generally positive view of psychological experimentation. The favorability of this view often diminished, however, as additional procedural information was presented. The ameliorative effect of debriefing information as measured by "willingness to participate" was weak, being contingent on the presence of desirable behavioral expectation information. In contrast, the ameliorative effect of debriefing as measured both by the percentage of definite refusals and "personal reaction" was substantially stronger. Indeed, regardless of the presence or nature of behavioral expectation information, debriefing raised "personal reactions" to stressful experiments to about the level associated with nonstressful experiments. Finally, on the "evaluation of experiments" measure, debriefing information eliminated the single stress effect detected; further, regardless of the presence or nature of behavioral expectation information, debriefing raised favorability of evaluation to that found in the rationale condition.

The differential effectiveness of debriefing as measured by "personal reaction" in contrast to "willingness to participate" is of interest. While debriefing information increased the favorability of the subjects' personal reactions to stressful experiments to about the level associated with nonstressful experiments, such information did not concomitantly increase willingness to participate. In terms of the respective scales, debriefed subjects indicated on the average that they would feel about the same after stressful experiments as they did before them, while also indicating that "maybe" they would be willing to participate in such studies.

Thus, within the current study, the fact that the subjects on the average believed they would not leave stressful experiments with greater anxiety or lower self-esteem than when they first arrived was not a sufficient condition for eliciting a high level of consent. The ethical requirement that the experimenter must not only remove stressful after-effects, but must additionally provide the subject with positive gain, would therefore appear judicious.

#### THE MILGRAM OBEDIENCE EXPERIMENT

While responses *across* either stressful or nonstressful experiments were assessed in the preceding analyses, individual reactions to the Milgram experiment are of particular interest. These may be compared with consent-related responses obtained in an experiment conducted by Ring, Wallston, and Corey (1970) and, thus, the predictive validity of the current role-play-derived consent estimates can be roughly assessed.

Ring et al. (1970) essentially replicated Milgram's (1963) procedure with the intent of empirically resolving some of the ethical questions surrounding the paradigm. After actually participating in the replication, the subjects completed a questionnaire in which their candid reactions to the experiment were solicited. Some of the subjects were given debriefing information before filling out the questionnaire; others were not. The questionnaire was presented to the subject as an attempt to determine "whether any experiments in which you've participated in any way violate the rights of subjects as stated in the Ethical Standards Code of the American Psychological Association [Ring et al., 1970, p. 74]."

On a dichotomous consent-related response, 4% of the Ring et al. subjects who had received debriefing information indicated that they regretted they had participated in the experiment; on a related dependent measure, 4% of the debriefed subjects indicated the experiment should not be permitted to continue. The corresponding percentages for subjects who had *not* received debriefing information were 43% and 57%, or, on the average, 50%. Debriefing, thus, had a substantial ameliorative effect on subjects who actually participated in this replication of the Milgram paradigm.

When the current role-play responses for the Milgram experiment alone are examined via one-way analyses of variance, a highly significant in-

TABLE 5

## Role-Playing Responses Obtained for Milgram Study

Variable	Information level						
	Rationale ( <i>n</i> = 19)	Procedure ( <i>n</i> = 18)	Desirable behavioral expectation ( <i>n</i> = 17)	Undesirable behavioral expectation ( <i>n</i> = 22)	Neutral debrief ( <i>n</i> = 8)	Desirable debrief ( <i>n</i> = 11)	Undesirable debrief ( <i>n</i> = 11)
Would you have chosen to participate? <sup>a</sup>	3.69 <sub>b</sub>	1.83 <sub>a</sub>	2.00 <sub>a</sub>	1.68 <sub>a</sub>	3.38 <sub>b</sub>	3.00 <sub>b</sub>	3.19 <sub>b</sub>
Percentage checking "definitely not"	5	61	53	77	12	9	9
Personal reaction following experiment <sup>b</sup>	11.16 <sub>b</sub>	6.12 <sub>a</sub>	6.06 <sub>a</sub>	4.68 <sub>a</sub>	9.38 <sub>b</sub>	9.64 <sub>b</sub>	9.55 <sub>b</sub>
Evaluation of experiment <sup>c</sup>	16.26 <sub>c</sub>	8.61 <sub>a</sub>	10.53 <sub>ab</sub>	6.96 <sub>a</sub>	13.13 <sub>bc</sub>	12.64 <sub>bc</sub>	12.91 <sub>bc</sub>

Note. Within a particular row, means with common subscripts do not differ, at the .05 level, as determined by the Duncan multiple-range test.

<sup>a</sup> 1 ≈ "definitely not," 2 ≈ "probably not," 3 ≈ "maybe," 4 ≈ "probably," 5 ≈ "definitely."

<sup>b</sup> 2 ≈ "sadder/less satisfied than before," 10 = scale midpoint, 18 ≈ "happier/more satisfied than before."

<sup>c</sup> 2 ≈ "not worthwhile/not valuable," 10 = scale midpoint, 18 ≈ "worthwhile/valuable."

formation effect on each of the continuous response measures is revealed. The relevant data are illustrated in Table 5. Furthermore, the effect was consistent across these measures. Favorability of response was highest in the rationale condition, was significantly reduced with the addition of procedural information (regardless of behavioral expectation information), and was significantly increased with the addition of debriefing information (again regardless of behavioral expectation information) such that the debrief cells did not differ significantly from the rationale condition. Thus, debriefing information, as in the Ring et al. replication, had a reliably ameliorative effect on consent-related responses.

A comparison of role-play-derived consent rates with rates derived upon the Ring et al. replication adds further support to the predictive validity of the role-play procedure. As before, our subjects' responses to the participation question were dichotomized (see Table 5).<sup>5</sup> Of the subjects exposed to the cover story, procedure, and debriefing information in the present study (neutral debrief), 12%

indicated they would "definitely not" wish to participate; the corresponding percentage for the subjects who did *not* receive debriefing information was 61%.<sup>6</sup> A comparison of these role-playing consent estimates with consent-related rates obtained from actual participants in the Ring et al. study indicates that our estimates are not badly off the mark. These estimates (12% for debriefed subjects and 61% for nondebriefed subjects) tend to be on the conservative side such that they are *overestimates* of the percentage of subjects who actually reported participation undesirable in the Ring et al. experiment (4% for debriefed subjects and 50% for nondebriefed subjects). This conservative bias may be due to the possibility that the Milgram (1963) procedure was more severe than the Ring et al. (1970, p. 77) procedure.

Thus, not only does the proposed role-playing procedure possess face validity, but, more importantly, the correspondence between the Ring et al. findings and our role-playing estimates suggests the procedure possesses predictive validity as well. If experiments routinely included an ethics-related questionnaire which could be anonymously completed by the subjects, as in the Ring et al. study, the validity and generality of role-playing-derived consent estimates could be better established. Of equal importance, such questionnaire feedback

<sup>5</sup> The critical reader may wonder why we persisted in dichotomizing the original consent measure *between* "definitely not" and "probably not." On a priori grounds, both responses would appear indicative of consent refusal and might be so scored. Such scoring yields consent refusal rates of about 16% for the nonstressful experiments in the rationale and/or procedure conditions. There would, however, seem little for subjects to be genuinely upset about since the experiments appear innocuous in these conditions. Thus, such scoring was judged to be empirically inappropriate (see Footnote 4).

<sup>6</sup> These role-playing-derived estimates were based only on the procedure and neutral debrief conditions since behavioral expectation information is typically unavailable to the experimenter.

would provide a stimulus for procedural alteration, should this be deemed necessary.

#### BACKGROUND MEASURES

To explore the relationship between the dispositional variables collected (self-esteem, need for social approval, attitudes toward psychological experimentation, etc.) and consent (scaled as a continuous variable), a number of correlational analyses were conducted.

To correct for the effects of the manipulated stress variable, correlations between the dispositional variables and the consent variable were run within each stress condition. That is, one set of correlations was run so as to predict the average consent response for the four stressful experiments, and a second set of correlations was run in an attempt to predict the average consent response for the two nonstressful experiments. In each of these analyses, the average consent response was expressed as a deviation about its associated information cell mean so as to correct for the simple effects of information. In analyzing the consent responses to the stressful experiments, for example, a rationale condition subject's average response to the four stressful experiments was computed and was then expressed as a deviation from the mean response in the rationale condition.

Five dispositional variables accounted for 20% of the consent response variance ( $F = 6.31$ ,  $df = 4/101$ ,  $p < .0005$ ) for stressful experiments and 2% of the variance ( $F = .52$ ,  $df = 4/101$ ,  $ns$ ) for nonstressful experiments:

(a) *Attitude toward psychological experimentation* was measured as the sum of responses to five 9-point scales which were, respectively, bounded by the descriptors: interesting–dull, worthwhile–not worthwhile, pleasant–unpleasant, valuable to science–not valuable to science, and soothing–threatening. The more favorable a subject's attitude toward experimentation, the greater his willingness to participate in stressful experiments ( $r = .38$ ,  $df = 104$ ,  $p = .0001$ ; nonstressful  $r = .05$ ,  $ns$ ).

(b) The more favorable a subject's *anticipated personal reaction to experiments* as measured by the sum of responses to four 9-point scales (happier than before–less happy than before, more satisfied with yourself than before–less satisfied with yourself than before, more tense than before–more calm

than before, and more favorable toward psychology than before–less favorable toward psychology than before), the greater his willingness to participate in stressful experiments ( $r = .29$ ,  $df = 104$ ,  $p = .003$ ; nonstressful  $r = .07$ ,  $ns$ ).

(c) and (d) Similarly, the more favorable a subject's *attitude toward psychology* as measured by a 9-point scale (extremely favorable–extremely unfavorable) and the more favorable his *attitude toward psychologists* as defined as the sum of responses to nine evaluative dimensions (e.g., trustworthy–untrustworthy, interesting–dull), the greater his willingness to participate in stressful experiments ( $r = .29$ ,  $df = 104$ ,  $p = .003$ ;  $r = .14$ ,  $df = 104$ ,  $p = .15$ ); for nonstressful experiments these variables respectively correlated .11 and .04 with consent.

(e) In addition, the higher a subject's *self-esteem*, as measured by the Berger scale, the less he tended to be willing to participate in stressful experiments ( $r = -.12$ ,  $df = 104$ ,  $p = .22$ ; nonstressful  $r = .04$ ,  $ns$ ).

#### Summary

A procedure in which a sample of subjects from a potential population are asked to role play a proposed experimental procedure and to judge their willingness to participate in that experiment had they been fully informed was investigated as an alternative to current practices for determining whether consent can be anticipated. On the basis of the current study, the procedure appears to possess predictive validity. Validity could be more generally ascertained, however, if experimenters regularly administered a postexperimental, ethics-related, questionnaire to the subjects. Responses to such a questionnaire could additionally be submitted as partial fulfillment of an ethics requirement, as well as submitted to individuals challenging the ethicality of a study.

In determining whether informed consent would be likely, the procedure would appear to have the virtue of being less susceptible to error than would judgments made by the principal investigator or by advisory board members who usually differ from the subject population on a wide variety of characteristics which may be critical for determining consent. The procedure has the further advantage of being an inexpensive and quick device for screening a variety of procedures, all of which may be

equally acceptable on methodological grounds, without extensive pretesting and actual participation by subjects. It might be noted that the role-play procedure also may be used to evaluate the adequacy of various "payment" plans for experimental participation and to determine whether the proposed plan assures "freedom from coercion to participate" (cf. Section 5 of APA Ethical Standards, Cook et al., 1972, p. VIII) for a particular subject population as opposed to another subject population.

Adequacy of representation of the information provided sample subjects may be judged by an advisory board. Information concerning likely subject behavior along the socially desirable-undesirable dimension probably ought to be provided where it is known, as well as information concerning the hypotheses under test and the purpose of the experiment, since these factors affected level of consent within the current study.

One problem with the role-play consent procedure is that, paradoxically, while deception methodology is usually used in an attempt to induce "naive" reaction to the independent variables under investigation, the procedure may "contaminate" the subject population for this type of experiment. One possible way to avoid this effect may be to screen cross-university or cross-city, etc., choosing the sample from a population deemed to be as similar as possible to that proposed for actual use (e.g., two midwestern universities).

While use of the proposed procedure would not totally guarantee that subjects' rights and welfare would be protected, it might provide a data base that would reduce the necessity for scientists to deal with subjective, conflicting, and emotional judgments on this important problem, as is too frequently the case.

## REFERENCES

- BARBER, T. X., & HAHN, K. W. Physiological and subjective responses to pain producing stimulation under hypnotically suggested and waking-imagined "analgesia." *Journal of Abnormal and Social Psychology*, 1962, **65**, 411-418.
- BAUMRIND, D. Some thoughts on ethics of research: After reading Milgram's "Behavioral study of obedience." *American Psychologist*, 1964, **19**, 421-423.
- BAUMRIND, D. Principles of ethical conduct in the treatment of subjects: Reaction to the draft of the Committee on Ethical Standards in Psychological Research. *American Psychologist*, 1971, **26**, 887-896.
- BERGER, E. M. The relation between expressed acceptance of self and expressed acceptance of others. *Journal of Abnormal and Social Psychology*, 1952, **47**, 778-782.
- BERSCHIED, E., BOYE, D., & WALSTER, E. Retaliation as a means of restoring equity. *Journal of Personality and Social Psychology*, 1968, **10**, 370-376.
- COOK, S. W., HICKS, L. H., KIMBLE, G. A., MCGUIRE, W. J., SCHOGGEN, P. H., & SMITH, M. B. Ethical standards for research with human subjects. (Draft) *APA Monitor*, 1972, **3**(5), I-XIX.
- DEUTSCH, M., & KRAUSS, R. M. The effect of threat on interpersonal bargaining. *Journal of Abnormal and Social Psychology*, 1960, **61**, 181-189.
- FESTINGER, L., & MACCOBY, N. On resistance to persuasive communications. *Journal of Abnormal and Social Psychology*, 1964, **68**, 359-366.
- FREEDMAN, J. L. Role-playing: Psychology by consensus. *Journal of Personality and Social Psychology*, 1969, **13**, 107-114.
- MILGRAM, S. Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, 1963, **67**, 371-378.
- ORNE, M. T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 1964, **17**, 776-783.
- ORNE, M. T., & SCHEIBE, K. E. The contribution of non-deprivation factors in the production of sensory deprivation effects: The psychology of the "panic button." *Journal of Abnormal and Social Psychology*, 1964, **68**, 3-12.
- RING, K., WALLSTON, K., & COREY, M. Mode of debriefing as a factor affecting subjective reaction to a Milgram-type obedience experiment: An ethical inquiry. *Representative Research in Social Psychology*, 1970, **1**, 67-88.
- U.S. PUBLIC HEALTH SERVICE. *Protection of the individual as a research subject*. Washington, D.C.: U.S. Government Printing Office, 1969.