

On the Ethics of Psychological Research

BARRY R. SCHLENKER AND DONELSON R. FORSYTH

University of Florida

Three philosophical positions on the nature of ethics were examined and contrasted within the context of psychological research: (a) teleology, which involves the balancing of the costs and benefits associated with an action as a means of developing general ethical rules; (b) deontology, which involves the rational adherence to rigid, universal rules that hold irrespective of the situation or consequences; and (c) skepticism, which involves denying the ability to apply universal rules and asserting the individuality of moral codes. Two experiments investigated the degree to which judgments of the ethicality of psychological research are affected by the consequences of the research and judge's ethical ideology. Subjects judged experiments that investigated obedience to an authority as generally less moral and more threatening to the participants' dignity and welfare when the proportion of total obedience was high rather than low. The proportion of participants (10 vs 40%) who were described as "psychologically upset" by the research did not affect moral judgments. These results were obtained across two different obedience situations (Milgram's study and a Watergate study) and different subject populations (high school and college students, males and females). As suggested by philosophers, a judge's ethical ideology determined how the perceived benefits and costs of the research were correlated with moral judgments. Teleologists weighed scientific benefits heavily, deontologists weighed participants' costs heavily, and skeptics weighed both heavily.

Ethical issues in psychological research are intimately tied to more general moral positions held by psychologists and by the population at large. The answers to questions such as "Should some minimal amount of potential harm to subjects be tolerated in order to advance science?" and "Should subjects be deceived during experiments in order to obtain potentially important information?" depend not only upon one's view of science, but also upon one's general ethical philosophy. The present paper seeks to accomplish two parallel objectives. First, it places into an historical and philosophical perspective the major positions taken by

The present project was facilitated by the support of the Organizational Effectiveness Research Program, Office of Naval Research (Code 452), under Contract No. N00014-75-C-0901; NR 170-797. The authors wish to thank Patricia Schlenker, Carol McKinney, Nancy H. McGinnis, and Howard Goldman for their comments and assistance, and Mr. Ron Larder, Bette Ackerman, and Kathy Sheley for encouraging their students to participate as research consumers and subjects. Requests for reprints should be addressed to Barry R. Schlenker, Department of Psychology, University of Florida, Gainesville, Florida 32611.

psychologists prescribing how ethical questions *should* be answered. Second, it reports two experimental investigations of some of the situational and individual-difference factors that influence how individuals *do* make ethical judgments.

Disagreements over answers to ethical questions abound within the psychological profession. Yet, as Cook (1975, p. 68) recognized, "A scientific and professional discipline is based in part on a set of shared values." The lack of consensus and the high degree of value conflict that currently exist necessitate giving the ethical issues raised by psychological research "a high priority for the immediate future." Although it is impossible to gain complete consensus on any one ethical position, it is unfortunately true that some psychologists are not even sure of some of the fundamental ethical issues, much less the alternative ethical positions that have been suggested. There seems to be a general groping without direction and a vague uneasiness at the mere mention of ethical concerns in psychological research. Placing ethical issues into perspective seems particularly crucial at this time since a survey has indicated that very few psychologists have ever read works by philosophers (Kindler & Gergen, 1965), and virtually nothing has appeared in the psychological literature to tie the various positions to their philosophical origins. Although psychologists undoubtedly subscribe privately to enumerable ethical positions, three in particular have generated a great deal of public discussion: the position of the Ad hoc Committee on Ethical Standards in Psychological Research, whose views were adopted by the Council of Representatives of the American Psychological Association and are presented in the *Ethical Principles in the Conduct of Research with Human Participants* (1973); the position taken by Diana Baumrind (1964, 1971); and the position taken by Kenneth Gergen (1973). These divergent ethical positions correspond quite closely to what philosophers have termed teleology, deontology, and skepticism, respectively.

Teleology

The teleological position proposes that the ultimate judgment of the morality of an action or set of actions depends upon the *consequences* produced by it. One is ethically bound to act in a way that produces "good" consequences, with good being variously defined in terms of pleasure, happiness, self-realization, fulfillment, and/or demand. The position has a long history, being proffered by Socrates, Plato, Epicurus, and utilitarians such as Jeremy Bentham and John Stuart Mill. Bentham's dictum, that actions should produce the greatest good for the greatest number, epitomizes the approach. A teleological approach to research ethics advises that the potential benefits of the research (e.g., advancement of science, beneficial technological applications, advantages to

subjects) must be weighed against the potential costs (e.g., harm to subjects, detrimental technological applications).

The Ethical Principles in the Conduct of Research with Human Participants is fundamentally a teleological ethical stance. The *Ethical Principles* assert that "The general ethical question always is whether there is a negative effect upon the dignity and welfare of the participants that the importance of the research does not warrant. . . . The nearest that the principles in this document come to an immutable 'thou shalt' or 'thou shalt not' is in the insistence that the human participants emerge from their research experience unharmed—or at least that the risks are minimal, understood by the participants, and accepted as reasonable" (APA, Committee on Ethical Standards in Psychological Research, 1973, p. 11). From this guiding benefit/cost principle, other principles are derived to serve as rules for the conduct of research. These include obtaining informed consent; remaining open and honest with the participants; respecting the participants' freedom to decline participation; insuring the confidentiality of the participants' data; protecting the participants from physical and mental discomfort, harm, and danger; completely debriefing the participants; and removing any undesirable effects of the research. Of course, the guiding principle allows an investigator to make exceptions to each of these subordinate rules, and the *Ethical Principles* provide a number of suggestions that can be followed by the investigator who is attempting to make a decision about the ethicality of an exception. These suggestions include consultation with others, careful analysis of the possible benefits (to the general public, science, or the participants' self-insight), and careful analysis of the prospective costs (violation of privacy, mental stress, etc.). The final decision as to whether or not to proceed with a study that violates one of the subordinate rules is left up to the researcher, who has personal responsibility for a study's ethical acceptability. However, this decision is still constrained by the guiding teleological principle.

Deontology

A deontological ethical philosophy rejects a rule or action's consequences as a basis for moral evaluation and appeals to natural law and rationality to determine ethical judgments. To the deontologist, acts are to be judged as moral or immoral through their comparison with some universal moral rule to which no exceptions can be made. Immanuel Kant, generally regarded as the foremost proponent of the deontological position, prescribed that one must "act only on that maxim which you can at the same time will to be a universal law" (cited in Frankena, 1973, p. 30). For Kant, this general universal principle or "categorical imperative would be that which represented an action as necessary of itself without reference to another end" (Kant, 1873/1973a, p. 75). Applied to

interpersonal conduct, Kant (1873/1973a, p. 82) deduces that one must "act to treat humanity, whether in thine own person or in that of any other, in every case as an end withal, never as a means only."¹

Deontologically, a moral principle can allow no exceptions, regardless of the consequences. Although Kant's ethical theory might appear simply to promote the performance of acts that produce "good" consequences, Kant held that rationality and natural law supplant the test of consequences. In his essay, "On the supposed right to tell lies from benevolent motives," Kant (1873/1973b) proposed that "to be truthful in all declarations is . . . a sacred unconditional command of reason, and not to be limited by any expediency" (p. 256), and that "all practical principles of justice must contain strict truths . . . since exceptions destroy the universality, on account of which alone they bear the name of principles" (p. 258). Kant argued, for example, that the universal maxim "always keep your promises" was a command of reason that should be acted on irrespective of the consequences. In "proving" his argument, he stated that if people acted on the opposite universal maxim (i.e., "only keep your promises when it is to your advantage"), the nature of that principle would negate the concept of a promise and hence "destroy itself" (see Frankena, 1973). Because the maxim "always keep your promises" does not contain such an internal inconsistency, Kant concluded that it must be a universal moral principle. Thus a lie, however benevolent, is intrinsically immoral. In fact, Kant asserted that if asked, one must truthfully tell a potential murderer the whereabouts of the prospective victim, since to do otherwise would be to break the universal rule. Kant does not insist that one must volunteer information that one would rather withhold, but every question must receive a truthful answer, and every statement must be truthful. For Kant, any form of deception in psychological research would be *a priori* immoral if it either provided subjects with untruthful information or treated them merely as means to a goal rather than as ends in themselves.

In her reactions to the *Ethical Principles* of the American Psychological Association, Baumrind (1971) echoed Kant's deontological stance. She contended that because the *Ethical Principles* focus on a "risk/benefit ratio" assessment of morality, they only serve to justify exceptions to fundamental moral principles and thereby violate the basic rules of ethical justice. "*Fundamental moral principles of reciprocity and justice*

¹ There are two subcategories of deontology, act-deontology and rule-deontology. In act-deontology, any action which falls into a particular category must be judged as good or bad irrespective of the circumstances or consequences. In rule-deontology, actions are considered good or bad on the basis of their adherence to general rules that apply irrespective of the circumstances or consequences. Rule-deontology has been considered the only viable form of deontology, since categorizing acts invariably assumes a guiding rule into which the categorization fits (cf. Frankena, 1973). Kant was a rule-deontologist.

are violated when the research psychologist, using his position of trust, acts to deceive or degrade those whose extension of trust is granted on the basis of a contrary role expectation. . . . Scientific ends, however laudable they may be, do not themselves justify the use of means that in ordinary transactions would be regarded as reprehensible. . . . The risk/benefit ratio justifies the sacrifice of the welfare of the subjects in the name of science, thus creating moral dilemmas for the investigator, and, as such, is not moral" (Baumrind, 1971, p. 890, italics added).

It should be noted that "fundamental moral principles" are violated by deception, as Baumrind contends, only so long as one subscribes to the deontological school of thought. Whenever a teleologist and a deontologist discuss morals, there will be an inevitable value conflict. The fundamental moral guide for the teleologist is the test of consequences, a consideration that the deontologist cannot abide. For the former, exceptions to subordinate rules are an inevitable though unwelcomed occurrence, while for the latter, exceptions are anathema. As with so many other differences between Kant and the British empiricists, rationality forms the code for the deontologist, while empirical consequences form the code for the teleologist.

Difficulties are inherent in both approaches, becoming particularly evident when applied to the ethics of psychological research. For example, within a teleological perspective, how are the rewards and risks to be quantified and balanced by the researcher interested in calculating the ethicality of a proposed project? If the consequences of the project are to provide the test of morality, how can the researcher take into account all of the consequences prior to the research since some may be unforeseeable? Further, a teleological position implies the existence of universal utility scales through which the precise values of a given set of consequences could be absolutely derived; psychometric methods have yet to produce a satisfactory scale that could be used in this context. In regard to the deontological perspective, the use of rigid and exceptionless principles appears to some to be both idealistic and impossible. Should one really abandon an important research area simply because the research involves deception? Many people would argue that abandoning potentially important research for such an "inconsequential" reason is itself unethical.

Skepticism

The diversity of opinions concerning moral principles and their exceptions led to the development of ethical skepticism. Although the several varieties that can be subsumed under this heading are often at odds with one another, they all share the similar assumption that violate moral codes cannot be formulated. In general, this approach, which includes emotivism, cultural relativism, and ethical egoism, recog-

nizes that there are many moral points of view, and all seek in one way or another to criticize those who attempt to present specific ethical principles.²

Emotivism grew out of logical positivism, which holds that the meaning of anything can only be decided on the basis of how one would go about measuring it and verifying it. Emotivists argue that moral judgments amount to nothing more than commands, appeals, or statements when they are stripped of excess meanings that cannot be seen, touched, heard, or otherwise sensed. Because they are simply commands, appeals, or statements, they cannot of themselves be right or wrong, moral or immoral. The emotivist views moral questions as ones without any substance.

Cultural relativism is a skeptical position that emerged from anthropological findings of drastic moral code differences between societies. According to the position, all moral standards are relative to the society in which they occur, so one cannot determine what is ultimately right or wrong. The best one can do is to show that an action or set of actions is consistent or inconsistent with the predominant patterns within a particular society. Blanshard (1966, p. 5) suggests that David Hume was the first cultural relativist: "Hume held that to approve of conduct is to have a certain feeling about it, a feeling caused by its perceived tendency to increase the happiness of people with whom we can sympathize; and to pronounce it right is to say that in our society people generally feel that way about it."

Adherents to another form of skepticism, *ethical egoism*, claim that no moral standards can be considered valid except in reference to one's own behavior. The only moral pronouncement possible is that one should act according to what one feels is right, and not act in a way that one feels is wrong. When universally applied, the pronouncement becomes "everyone should always act so as to promote his own interest," a position adopted by Hobbes as an enlightened egoist (Davis, 1973). At one level, ethical egoism is similar to teleology, since egoists typically allow consequences to serve as the basis for determining right or wrong. The egoist, however, avoids the problem of developing transpersonal utility scales, since only his or her own value judgments are crucial. Each person is allowed to determine idiosyncratically the weights and values

² Situation ethics (Fletcher, 1966) has also been considered a type of ethical skepticism. Fletcher argues that there is one universal guide, "to love God and your neighbor," into which all actions must be "fitted" according to the situation to yield "contextual appropriateness." Even though situation ethics allows situational relativity, the position still espouses the fundamental and invariant rule noted above, and some philosophers believe Fletcher merely skirts the issue of invariant rules (e.g., Davis, 1973).

of particular outcomes, and individual moral codes may therefore differ drastically.

Kenneth Gergen's (1973) comments on the *Ethical Principles* adopted by the American Psychological Association fall under the rubric of ethical egoism. Gergen disapproves of attempts to develop codes of moral principles, believing that such codes are detrimental to both the profession and to the accumulation of knowledge. "An appeal to abstract moral principles, cut away from their precise consequences in each situation, sets the stage for all manner of tyranny, from preventing the accumulation of knowledge to mass murder. . . . Absolute moral values are absolutely corrupt" (Gergen, 1973, p. 908). As an alternative to a systematic set of ethical principles, Gergen proposes that psychologists should conduct research to determine the actual consequences of various investigatory procedures for the participants. "Advisory statements" can then be accumulated that describe how various research procedures affect the participants' outlooks on research, the amount of anxiety, stress, and harm they experience, and so on. "Apprised of the potential consequences of his actions, the researcher is left free to deliberate the ethical implications of his actions. . . . Advisory statements would encourage independent ethical decision-making in a context of factual enlightenment. . . . What is being advocated here is that we accept the widely divergent ethical values within society on a par with our own" (Gergen, 1973, p. 908).

William James (1891/1973) was an ethical skeptic and similarly discussed the difficulty of judging the morality of an action prior to the availability of facts about it. In developing his position, James looked to consequences as the ultimate test of morality. However, James was critical of the views of the utilitarians because he felt that they excluded too much of what people regard as morally binding or good, and he objected to Kant's views for much the same reason. In order to include the entire panorama of values that people might hold, James (1891/1973, p. 154) proposed that "the essence of good is simply to satisfy demand," and defined demand as "anything under the sun," meaning whatever values people feel must be satisfied. His position has been termed *value pluralism* because of the allowed inclusion of all values. Whether or not any particular system of values successfully satisfies demands, though, cannot be decided in advance but instead must be determined by its consequences after it has actually been put into practice. "These experiments [the implementation of value systems] are to be judged, not *a priori*, but by actual finding, after the fact of their making, how much more outcry or how much appeasement comes about. . . . Everywhere the ethical philosopher must wait on facts" (James, 1891/1973, p. 157). For James, it would be difficult if not impossible to judge the morality of any

new research procedure that could potentially "satisfy demands" until information is available regarding its actual consequences.

Summarizing the Differences

Two major distinctions between the three approaches to moral philosophy are notable and relevant to the present research. The first concerns the willingness to proffer the existence of universal moral codes. Deontologists assert that universal ethical principles exist and must be followed without exception. Teleologists similarly insist that universal principles exist (based on a benefit/cost ratio), though they are willing to tolerate exceptions under special circumstances. (For example, a teleologist might insist that obtaining informed consent from research participants is a universally applicable ethical rule but would admit an exception if an experiment could not be otherwise conducted and risks to participants are judged as negligible.) Skeptics, on the other hand, deny the possibility of developing universal ethical rules.

Second, the positions differ in the degree to which they endorse idealistic versus pragmatic views. If an act fails to meet the standards of a universal rule, deontologists should condemn it regardless of the amount of harm or benefit produced by it. In this sense, deontologists can be described as idealistic. Teleologists are willing to tolerate negative consequences to the degree that positive consequences outweigh them and hence are more pragmatic. Skeptics should similarly be guided by consequences information, but there may be high variability across skeptical judges, with some evidencing more idealistic judgmental patterns than others. Thus, some skeptics might be as idealistic as deontologists and others as pragmatic as teleologists, but, in either case, a skeptic should differ from adherents to the other two positions by denying the applicability of universal moral rules.

RESEARCH ON ETHICS

The positions of both Gergen and the teleologists are compatible with the idea that we should perform research on the consequences of research. It should be kept clearly in mind, though, that moral questions are quite different from scientific ones. Science can only perform a descriptive function, not a prescriptive one. Ethics serve to establish the criteria used for moral judgments, while science can help determine whether or not something meets these criteria. As such, science can provide answers to ethically related questions such as those concerned with: (a) the means that can be used to obtain particular values or goals; for example, can school busing successfully achieve integration?; (b) the consequences of implementing specific values or goals; for example, what psychological, sociological, economic, and political changes accompany integration?;

(c) the attitudes and behaviors of people in relation to specific values and goals; for example, how do people feel and act toward integration?; and (d) the factors (social, developmental, etc.) that influence moral evaluations; for example, under what conditions will certain people evaluate integration as just or unjust? The question of whether a specific value or goal *is* moral or immoral is not a scientific question and falls squarely into the realm of moral philosophy.³

The present experiments examined how information about the consequences of a research procedure affect moral evaluations of it. Since experimentation on obedience to authority has generated ethical controversy, Milgram's (1963, 1965) investigations of people's willingness to harm another person at the command of an experimenter and West, Gunn, and Chernicky's (1975) investigation of obedience in committing a Watergate-like burglary were selected for evaluation by subjects. Subjects read a detailed description of one of those experiments and then evaluated its ethicality and information value. The consequences of each study were systematically manipulated by varying the proportion of participants who were described as completely obedient (high or low obedience) and the proportion who experienced extreme psychological upset during the study (high or low upset). In addition, subjects in Experiment 2 completed a moral opinion questionnaire. Factor analysis of responses to that questionnaire allowed subjects to be grouped according to their endorsement of a teleological, deontological, or skeptical position. Differences between these groupings in ethical judgments of the obedience studies could then be assessed.

The pioneering work in attribution theory done by Heider (1944, 1958), Jones and Davis (1965), and Kelley (1967, 1971) has recently been applied to the area of moral judgments (e.g., Kelley, 1971; Ross & Ditecco, 1975). An attributional approach to moral judgment is concerned with the manner in which individuals, based on their knowledge of an actor, an action, and its effects, make an inference regarding the actor's ethicality. Two processes are important in such analyses. First, the act itself must be evaluated as "good" or "bad." Second, the relationship between the act, its effects, and the actor must be determined; that is, is the actor "responsible" for the effects?

Attribution theorists have devoted relatively little attention to the first

³ Kohlberg's (1969) work on the six stages of moral development is perhaps the most frequently reference psychological exploration of moral issues. Kohlberg focuses on the question of why a person holds particular standards (i.e., because he will be punished unless he does, because society says it is right, or because his conscience says it is right) and not on the nature of the standards themselves. Philosophers, in contrast, are concerned primarily with the nature and ramifications of the standards. Although Kohlberg's schema is assumed to be independent of a person's specific moral beliefs, future work will be needed to test this assumption.

of these processes, evaluating acts. Following Heider, Ross and Ditecco (1975, p. 92) assume that moral standards derive from "oughts," which are "impersonal standards that indicate what behaviors are appropriate in a particular situation." The limitation of such an analysis is that it explicitly assumes that *ought* demands apply cross-situationally and can be consensually validated by virtually all normal observers. Yet as our preceding discussion demonstrates, moral principles ("ought standards") are not viewed by everyone as being universal (cross-situational) and there is interpersonal disagreement as to what standards should be applied. It is hypothesized that deontologists (i.e., subjects who endorse the major tenets of that school on the ethical opinion questionnaire) should be concerned with whether or not a particular procedure violates a universal rule of reason and hence should not weight costs and benefits when arriving at an ethical evaluation of the obedience studies. Teleologists and most skeptics (as classified by their questionnaire responses) should be concerned with the balancing of costs and benefits in deriving their ethical evaluations. These hypotheses were tested in Experiment 2 by correlating ethical judgments with various perceived benefits and costs associated with the obedience studies. If people who endorse different philosophical positions do differ in the ways they process information when arriving at their ethical judgments, then differences in moral principles will have to be taken into account when applying attribution theory to the area of moral judgments.

Attribution theorists have been primarily concerned with assessments of an actor's responsibility. Generally, "the extent to which a person is viewed as responsible for a behavior is inversely related to the degree to which external factors are perceived to be determinants of the actions" (Ross & Ditecco, 1975, p. 92). The social desirability of an action, for example, affects the degree to which external restraints are seen as operating on the actor and hence the degree to which observers hold him responsible. Socially desirable actions have less information value than socially undesirable ones, with actors being held accountable for the latter more readily than the former (Jones & Davis, 1965; Ross & Ditecco, 1975). As will be seen, Milgram (1964) believes that the socially undesirable actions of his subjects affected attributions made about the ethicality of his studies.

Attribution theorists have also found that the magnitude of the consequences of an action affects observer's judgments of an actor's responsibility, even when the consequences are unforeseen and unintended. Walster (1966) found that subjects who read descriptions of an accident attributed more responsibility to the actor for causing the accident when the consequences were severe rather than mild. Subsequent investigations have both replicated (e.g., Chaikin & Darley, 1973; Harvey, Harris, & Barnes, 1975; Phares & Wilson, 1972) and failed to replicate (e.g.,

Shaver, 1970; Shaw & Skolnick, 1971) Walster's findings. Although numerous interpretations of these conflicting results are possible, at the very least the data indicate that under some conditions the consequences of an action will determine the degree to which a person is held accountable for it. From this conclusion, it is only a short step to propose that the consequences of an action (irrespective of foreseeability) will also affect moral evaluations of it. This point is particularly important in the case of research studies in which it is difficult to predict in advance the potential consequences to subjects. The severity-of-consequences phenomenon dovetails nicely with William James' views; an ethical philosopher must wait for facts about all of the consequences of an action before pronouncing judgment. Unanticipated consequences, particularly severe ones, can bring either moral condemnation or praise.

Milgram (1964) believes that attacks on his research program were precipitated by a combination of the severity-of-consequences phenomenon and the social undesirability effect. Under the guise of a learning experiment, Milgram ordered subjects to deliver increasingly severe electric shocks to a learner each time the latter made a mistake on a learning task. Contrary to expectations, a large proportion of subjects (65% in the initial studies) were totally obedient, delivering shocks through switches that were labeled "450 Volts, Danger: Severe Shock," and that about 10% of the subjects were so upset by the procedure that they seemed to be on the verge of a psychological breakdown. In response to Baumrind's (1964) criticism of the ethicality of his research, Milgram (1964, p. 849) raised the question, "Is not Baumrind's criticism based as much on the unanticipated findings as on the method? The findings were that some subjects performed in what appeared to be a shockingly immoral way. If, instead, every one of the subjects had broken off at 'slight shock,' or at the first sign of the learner's discomfort, the results would have been pleasant, reassuring, and who would protest?"⁴ Based on Milgram's arguments and the findings of attribution theorists, it was predicted that the more negative the consequences of the experiments, the greater the moral condemnation of the study.

METHOD

Subjects

Subjects in Experiment 1 were 61 male and 71 female high school seniors, while those in Experiment 2 were 87 male and 93 female college undergraduates who participated during the first week of an introductory psychology course. No subject had previously heard about or read the Milgram or West et al. obedience studies. (The studies were

⁴ At a prescriptive level, Milgram's question is irrelevant. Whether people do or do not make moral judgments by taking into account the consequences of an experiment does not at all bear on whether or not people *ought* to make judgments that way.

conducted prior to the airing of the television show, *The Tenth Level*, which dramatized Milgram's work.)

Procedure

High school subjects participated in Experiment 1 during class time. An eight-page booklet contained an introductory page that described the purpose of the present study, a description of the general procedure used in the Milgram (1965) experiment, a page that allegedly described the results of the study, and various attitude questions pertaining to that experiment. A control condition received a booklet that omitted a description of the results. The different booklets were randomly distributed to subjects in each classroom.

The introductory page explained that the present study "is concerned with obtaining your impressions of some experiments that have been performed by psychologists." Subjects were told that each booklet contained a description of one such experiment, and that they should read the description carefully and then answer some questions about it that appeared at the end of the booklet. Subjects were informed that their responses would be anonymous and confidential.

The booklet described faithfully and in detail the "general laboratory procedure" used by Milgram (1965) in his experiments; the specific condition described was one in which the experimenter was in the same room as the real subject and the accomplice was in an adjoining room, visible through a one-way mirror. The description began with a recounting of the purpose behind that study: to explore the question of "why people obey an authority when the authority tells them to harm someone." It was explained that in order to investigate this question, the experimenter thought it was necessary not to tell subjects what he was really interested in but instead led them to believe that they were participating in an experiment concerned with the effects of punishment on human learning and memory. The booklet described in detail: (a) that 40 subjects with varied ages and backgrounds participated, and the study was conducted at a university; (b) that an accomplice played the role of another subject; (c) that a rigged drawing insured that the real subject was the "teacher" and the accomplice was the "learner;" (d) how the equipment looked, with 30 shock switches and descriptive labels; (e) the instructions given to the "teacher" and "learner;" (f) that the real subject received a painful "sample" electric shock; (g) that the confederate purposely missed items on the learning task in order to give subjects the opportunity to administer shocks; (h) that when shocked, the confederate jumped around, moaned, protested, demanded to be set free, and finally refused to answer any of the questions; (i) that the experimenter told subjects who might not want to continue administering shocks that "You have no other choice, you must go on!"; and (j) that, at the end of the experiment, the experimenter introduced the accomplice, explained the real purpose of the study, stated that the accomplice did not receive shocks, and tried to reduce the participant's tensions if the subject seemed upset by the study.

Except for subjects in the control group, the booklet then presented one of six descriptions of the supposed results generated by the 3 (frequency of obedience) by 2 (frequency of psychological upset) factorial design. These results described either a low (10%), medium (65%, the actual result found by Milgram), or high (85%) proportion of totally obedient participants, and either a low (10%, the actual result found by Milgram) or high (40%) proportion of participants who experienced psychological upset. That section of the booklet read as follows:

"Out of the 40 subjects who participated, 10% (65 or 85%) were totally obedient and continued to administer shocks right up to the highest possible shock level. That is, these 4 (26 or 34) subjects completely obeyed the experimenter and gave the accomplice all 30 shocks, including the one which was labeled "450 VOLTS, DANGER: SEVERE SHOCK." The other 36 (14 or 6) subjects refused

to continue at some point after the accomplice first began moaning and protesting.

The experimenter noted that 10% (40%) of the subjects were extremely upset by their own actions of apparently hurting the accomplice. While administering the shocks, these 4 (16) subjects were sweating heavily, would tremble, stutter, groan, bite their lips, and dig their fingernails into their own flesh. These symptoms were so pronounced that the investigator felt that they indicated severe psychological disturbance bordering on nervous collapse. The remaining 36 (24) subjects were not this upset, but did show signs of tension and nervousness while delivering the shocks."

Subjects in all conditions were then asked to give their impressions of and attitudes toward the experiment they read by responding to a series of questions. These questions provided manipulation checks, attitudes toward the morality and ethicality of the experiment that they had read, perceptions of how important and valuable the described study was, and evaluations of the experimenter who conducted the experiment. Each of the major items was followed by a 15-point scale with labeled end points. Evaluations of the experimenter consisted of 20 15-point bipolar adjective scales.

Experiment 2 was virtually identical to Experiment 1 in procedure and was conducted to assess the generalizability of the results of Experiment 1 to a different population and a different experimental scenario. Experiment 2 differed from Experiment 1 in three ways. First, college students served as subjects in Experiment 2, allowing a possible assessment of the effects of hedonic relevance (Jones & Davis, 1965) on moral judgments. Unlike the high school subjects, for whom research was hedonically irrelevant, the college subjects knew that they would have to participate in at least 3 hours worth of experiments during the remainder of their psychology course. Second, after completing the experimental booklet, subjects also responded to a questionnaire designed to assess their agreement with certain tenets suggested by each of the various schools of moral philosophy. Third, subjects read either a description of the Milgram study used in Experiment 1 or a description of the procedures used by West et al. (1975) in the government sponsorship/immunity condition of their obedience study. Although both the Milgram and West et al. studies involved obedience, participants in the former were told to physically harm another person by administering electric shocks (but didn't agree to break any laws), while participants in the latter were asked to commit a burglary (but didn't physically harm another person). Thus, though conceptually similar, the different elements of the two studies provide a fair assessment of generalizability.

The Watergate booklet noted that the psychologist who conducted the study had contrived an experiment designed to investigate why people "obey an authority when the authority tells them to do something illegal." The description explained that people might not behave normally if the study were done in a laboratory or if they were told about its actual purpose, so a false "cover story" was made up that provided participants with a justification for the situation. The booklet described in detail: (a) the fact that criminology students were approached by the experimenter-detective and a subsequent meeting was arranged; (b) that subjects were told at the meeting that they would be discussing plans to burglarize a local advertising firm to obtain information about income tax evasion that was needed by the Internal Revenue Service; (c) that all participants in the theft would receive immunity from prosecution if apprehended; (d) the details of the crime; (e) the fact that the subjects participated in the planning phase of the crime and interchanged questions with the detective; (f) the request made by the detective for attendance at the "final" planning meeting; (g) the fact that the detective attempted to persuade participants who balked at attending the "final" meeting and committing the burglary; and (h) that at the close of the experiment the detective explained the reasons for the situation.

Except for control subjects, the booklet then presented one of four descriptions of the

supposed results using the same format as the Milgram scenario. Subjects were informed that either 10 or 85% of the participants "were obedient and agreed to participate in the burglary." It was also pointed out that either 5 or 40% of the participants displayed the same type of "severe psychological disturbance bordering on nervous collapse" that was described in the Milgram scenario, including the specific symptoms.⁵ As with the Milgram scenario, where psychological upset was linked to obediently harming another person as opposed to disobeying the experimenter, it was made clear that the symptoms were the result of the fact that the upset participants "were bothered by their own actions in the situation and that the active participation in the planning of a burglary that might be immoral threatened their self-esteem." The moderate obedience condition was deleted in Experiment 2. All percentages were identical for both the Milgram and Watergate scenarios except for the low upset condition. As will be seen, this slight difference between the low upset conditions in the two scenarios did not produce any effects. Thus, Experiment 2 was a 2 (Watergate or Milgram scenario) by 2 (low or high frequency of obedience) by 2 (low or high frequency of upset) factorial which included a no-results-presented control group for each of the two scenarios.

Ethical Positions Questionnaire

Subjects in Experiment 2 answered a 68-item questionnaire designed to assess individual differences in ethical positions. Subjects indicated the extent of their agreement or disagreement (on nine-point scales) with the first 50 items. Many of these items were selected to tap the major dimensions of ethical concern discussed by adherents to the three moral philosophies presented in the introduction. The items ran the gamut from ones concerned with the feasibility of universal ethical codes to ones concerned with deception, harm to research participants, and the ability of science to solve the world's problems. On the last 18 items subjects indicated the degree of importance (from "not at all important" to "primary importance") they felt should be attached to specific considerations when determining whether or not a psychological study should be conducted. Items included such concerns as the confidentiality of data, the enjoyableness of a study for participants, and the amount of harm that could occur.

RESULTS

Analysis of the Ethical Positions Questionnaire

Items from the ethical positions questionnaire were grouped into coherent categories using factor analysis with principal axes solutions and orthogonal varimax rotations. The first two factors that emerged produced meaningful configurations of items that reflected two important dimensions. The first factor appears representative of an *idealism-pragmatism* dimension relevant to the benefits and costs of research. Examples of the 14 items that loaded heavily on this factor (greater than .40) are presented in Table 1. Subjects classified as idealists insisted that no harm, however small, was permissible in research, that people's welfare was crucial, and that it was of primary importance that a project might advance science. The pragmatists, on the other hand, felt that some degree of harm was permissible and that it was not of primary importance for a scientific ad-

⁵ West *et al.* (1975) did not report any psychological upset in their study; this aspect was added to the booklets to allow a systematic manipulation of consequences.

TABLE 1
EXAMPLES OF ITEMS USED FOR ETHICAL IDEOLOGY CLASSIFICATIONS

Factor I: IDEALISM-PRAGMATISM (Eigenvalue = 7.32)	Factor I Loading	Factor II Loading
"When one weighs the potential benefits from research against the potential harm to participants, it could lead to sacrificing the participants welfare and hence is wrong."	.43	-.20
"If a researcher can foresee any type of harm, no matter how small, physical or psychological, he or she should not conduct the study."	.44	-.22
"The dignity of people is the most important concern in any society."	.41	.13
"Scientific concerns sometimes justify potential harm to participants."	-.47	.17
"The amount of psychological harm that could potentially occur to participants."	.71	.01
"The degree to which the research might advance science."	.47	.23
Factor II: RULE-UNIVERSALITY (Eigenvalue = 4.16)		
"It is possible to develop rigid codes of ethics that can be applied without exception to all psychological research."	-.13	.51
"Lying to participants about the nature of a study is always wrong, irrespective of the type of study or the amount of information to be gained."	-.17	.51
"What is ethical varies from one situation and society to another."	-.20	-.41
"The application of a particular code of ethics depends entirely upon the particular study; what is appropriate in one study might be totally inappropriate in another."	-.13	-.50
"Rigidly codifying an ethical position that prevents certain types of research could stand in the way of scientific advancement and prevent the accumulation of knowledge."	.04	-.50

vance to follow from research; a balancing of benefits and costs seemed paramount for these persons.

The second factor represented a *rule-universality* dimension pertinent to the degree subjects' felt that universal, relatively rigid ethical codes could be developed. Examples of the 13 items that loaded on this factor are also presented in Table 1. Rule-universality appears indicative of the deontological and teleological schools, while rule-nonuniversality appears most indicative of a skeptical school.

Each subject's average scores on the items that loaded heavily (rotated factor loadings greater than .40) on each factor were computed. By crossing the two factors (by performing a mean split on each), a four-way classification was created (cell sizes for each of the four groups are contained in Table 4). People classified as idealists who are rule-universal closely fit the classical description of the deontologist. Pragmatists who are relatively rule-universal resemble teleologists (though a few teleologists might also be found in the pragmatism/rule-nonuniversal cell). All skeptics are rule-nonuniversal but also might be either pragmatists or idealists. Although most skeptics fall into the pragmatist/rule-nonuniversal cell, some, particularly those who hold high personal standards but don't insist on those standards being applied to everyone, could also classify themselves as idealists who are rule-nonuniversal.

The ability to form these classifications empirically is of importance. Factor analysis demonstrated that the major dimensions debated by ethical philosophers are salient to laymen and can be used to differentiate people. The often cerebral debates of ethical philosophers concerning these dimensions are justifiable when used for descriptive purposes and should receive greater attention by psychologists. We will shortly return to the use of these groupings to differentiate reactions to the scenarios.

Perceptions and Predictions of Obedience and Upset

The data from Experiment 1 were subjected to a 3 (obedience) by 2 (upset) by 2 (sex) analysis of variance while the data from Experiment 2 were subjected to a 2 (obedience) by 2 (upset) by 2 (scenario) by 2 (sex) analysis of variance. All analyses of variance were performed using a least-squares regression procedure that adjusted for all equal and lower order effects in a stepwise manner, thus taking into account unequal cell sizes. Multiple comparison tests were performed using Duncan's Multiple Range test unless otherwise noted.

The obedience and psychological upset manipulations were quite successful. Obedience main effects were obtained in both experiments across both scenarios on an item assessing how obedient participants were ($p < .001$). The similarity between the comparable means in both experiments was striking. (Means for E1 and E2, respectively, were: low obedience, 8.7, 8.5; high obedience, 12.9, 12.2; control, 9.5, 10.4.) Main effects of psychological upset were also obtained in both experiments across both scenarios on an item assessing the amount of psychological harm that occurred (E1, $p < .05$; E2, $p < .073$). (Means for E1 and E2, respectively, were: low upset, 10.0, 11.1; high upset, 11.1, 11.9; control, 9.1, 11.6).

Subjects were also asked on the questionnaire to write down the number of people (from 0 to 40) whom they would have predicted: (a) would

completely obey the experimenter and either administer all of the shocks (Milgram scenario) or agree to take part in a burglary (Watergate scenario), and (b) would be extremely upset and disturbed by the experiment and experience psychological harm. Again, the manipulations affected the appropriate items ($ps < .02$), demonstrating both the effectiveness of the manipulations and the fact that subjects retrospectively bias what they believe they would have predicted after discovering the "actual" results of a study.

No other main effects or interactions between any of the independent variables were obtained on any of the manipulation check items except for a consistent main effect of scenario in Experiment 2. Subjects stated that participants in the Watergate as compared to Milgram scenarios were less upset ($p < .03$) and less obedient ($p < .01$), suggesting that the Watergate study was viewed as less coercive.

The obedience and upset manipulations were perceived as separate dimensions, since the effects of each were confined to their own manipulation check item. Also, predictions of obedience and harm were not significantly correlated either overall (E1: $r = -.16$, $n = 133$; E2: $r = -.14$, $n = 175$) or within the control conditions only (E1: $r = -.12$, $n = 20$; E2: $r = -.18$, $n = 36$). Subjects did not assume that obedience by itself was an indicator of harm; as will be seen shortly, some subjects may even have equated obedience with helping the experimenter.

Morality and Ethicality of the Experiments

Since the phrase has been frequently employed in discussions of research ethics, subjects were asked to judge, "How much do you feel the experiment threatened the dignity and welfare of the subjects?" A main effect of obedience was found on responses to the item in both experiments (see Table 2 for means and significance levels). Specifically, perceived threat to the participants' dignity and welfare was directly related to the proportion of reported obedience. However, no other main or interactive effects on responses to this item were found in either experiment. It seems unlikely that subjects believed that obedience was directly harming the participants, since predicted obedience and predicted harm were not correlated with one another, and the manipulation of psychological harm did not affect judgments of perceived threat. Instead, subjects were probably reacting to either the social undesirability of the behavior (as suggested by Milgram) or to the belief that compliance would occur only under extreme pressure by the experimenter, thus robbing the participants of the exercise of free choice (though not necessarily harming them).

Subjects were also asked to rate how ethical and moral they felt the experiment was. Although the means for the experimental obedience con-

TABLE 2
RATINGS OF MORALITY AND ETHICALITY OF EXPERIMENTS

	Obedience means				Significance
	High	Medium	Low	Control	
Item I:					
Threat to Dignity and Welfare					
Experiment 1	9.7 ^a	9.7 ^a	7.5 ^b	8.1 ^{ab}	$F(2,101) = 5.28, p < .007$
Experiment 2	9.4 ^a	—	7.9 ^b	9.1 ^{ab}	$F(1,126) = 6.01, p < .02$
Item II:					
Ethical and Moral					
Experiment 1	7.2 ^a	8.4 ^{ab}	8.5 ^{ab}	9.5 ^b	$tD(4,129) = 2.37, p = .05$
Experiment 2	8.2 ^{ab}	—	8.6 ^a	6.8 ^b	$tD(3,172) = 2.62, p = .05$

Note: Means within each row without a common superscript differ by at least $p = .05$. For Item I, higher means indicate greater threat; for Item II, higher means indicate greater perceived ethicality and morality.

ditions were in the predicted direction in both experiments, they did not differ significantly from one another. However, comparisons using Dunnett's procedure for contrasting treatment groups with a control condition revealed that control subjects in Experiment 1 did feel that the experiment was more ethical and moral than subjects in the high obedience condition (see Table 2). In Experiment 2, comparisons with the control condition revealed that subjects in the low obedience condition rated the experiments as more ethical and moral than did subjects in the control conditions (see Table 2). The means for the comparable conditions in both experiments are very similar except for the control groups; the college students in Experiment 2 rated the control scenarios as less ethical ($p < .05$) than the high school students in Experiment 1. A major difference between these control conditions that might account for the above effect was in their views of the relationship between obedience and morality. As suggested above, once subjects are informed of the often high and perhaps surprising levels of obedience associated with the studies, they may have viewed obedience as either socially undesirable or as reflecting tremendous pressure by the experimenter. But in the absence of explicit information about obedience, observers might assume that participants would obey only to the degree that they felt the study was moral and ethical. Such a positive correlation between perceived obedience and judgments of morality was found in the control condition of Experiment 1 ($r = .55, p < .01, n = 20$). However, no significant relationship was obtained between these variables within the control conditions of Experiment 2 ($r = -.12, ns, n = 36$); these correlations differ significantly from one another ($p < .01$). Thus, the high school control subjects equated

obedience with ethicality, while the college control subjects did not. Such a difference would account for the more positive ratings of morality by the high school control subjects. The college control subjects' failure to associate these variables could be due to the greater hedonic relevance of the research to these subjects; as a result of the relevance, they may have been motivated to dissociate obedience from moral considerations.

The results on the threat and morality items provide support for Milgram's contention that the behaviors of the participants in an experiment will affect the moral judgments made about it. Subjects in the experimental conditions in both experiments and across both scenarios judged the research to be more threatening and less moral the greater the obedience; as Milgram asserted, the social undesirability of the participants' behaviors may have reflected negatively upon the ethics of the study. Another possibility is that obedience was viewed as an indication of the amount of pressure the experimenter placed upon the participants. That is, observers, once aware of a study's results and participants' psychological reactions, may have taken obedience to be an indication of the degree to which the experimenter pressured participants to behave against their wills, thus making the study unethical. However, the fact that perceived obedience was not significantly correlated with judgments of the experimenter as "strong-weak," "nonaggressive-aggressive," or "active-passive" argues against this latter interpretation.

No between-conditions differences (except for a scenario main effect to be presented shortly) were found on an item asking "Do you feel that the experimenter could have foreseen any possible psychological harm to the subjects?" Subjects felt that harm clearly could have been foreseen (Experiment 1: $M = 10.2$; Experiment 2: $M = 9.7$). The relationship between harm and moral judgments will be considered in terms of within-conditions differences shortly.

Informational Value of the Experiment

In Experiment 1, an obedience by psychological upset interaction was obtained on an item which asked, "How much do you feel was learned from the study in terms of new and useful information?"; $F(2,101) = 3.24, p < .05$. As can be seen in Table 3 obedience was negatively related to how much was learned when upset was low. However, under high upset conditions, obedience was positively related to the amount learned. Comparison tests indicated that subjects in both the low-obedience/low-upset condition and the high-obedience/high-upset condition felt that significantly more was learned than did subjects in the control condition.

An obedience by psychological upset interaction was also found on this item in Experiment 2; $F(1,126) = 7.80, p < .01$. As also shown in Table 3,

TABLE 3

MEANS FOR THE OBEDIENCE BY UPSET INTERACTIONS ON RATINGS OF INFORMATION VALUE

	Experiment 1		Experiment 2	
	Low upset	High upset	Low upset	High upset
Low obedience	10.8	9.3	8.9	8.5
Medium obedience	9.1	10.1	—	—
High obedience	9.2	11.2	10.7	7.9
Control	8.2		8.4	

Note: Higher means indicate greater information value.

the pattern of this interaction differs greatly from that obtained in Experiment 1. The college students felt that the experiment was most informative and valuable when few participants were upset and obedience was high. Analyses of related measures (e.g., the degree to which the research should have been done, how scientifically valuable the study is, and how much the experimenter was concerned with contributing something important to science) yielded patterns identical to those presented on the "information" item in each experiment; detailed presentation of these measures is omitted for the sake of brevity.

One plausible explanation of the differences between Experiment 1 and 2 pertains to the greater hedonic relevance of the scenarios for the college than for high school students. An ideal experiment from a prospective subject's point of view might be one in which the researcher's hypotheses are supported and the participants emerge from the research with little or no psychological harm. Cooperating with the experimenter and supporting his hypotheses has been considered a major concern for subjects (Orne, 1962), and the descriptions of both scenarios pointed out that the researchers were examining obedience and presumably would appreciate some compliance. Thus, the most valuable research was judged as that which involved high obedience but little psychological harm. As uninvolved observers, the high school students might have been more likely to perceive research to be more valuable when the participants behave in ways that are opposite that which would be expected by the reward-cost structure of the situation (Jones & Davis, 1965). Under high upset conditions, the costs for complying with the experimenter's demands appear to be quite high; while under low upset conditions, the costs for complying appear to be relatively low. More was learned to the extent that the participants went against this structure and complied despite high costs (high-upset/high-obedience) or resisted complying even though the costs for complying were relatively low (low-upset/low-obedience condition).

Sex and Scenario Differences

Females reacted more negatively than males to the scenarios in both experiments, believing that the research was a greater threat to the dignity and welfare of the participants and that the experimenter was more cruel, unethical, and bad ($ps < .05$). These results support the stereotypes of males as more scientifically oriented and females as more humanistically oriented.

No significant interactions involving scenario occurred in analyses of any of the major dependent variables. Subjects did have a more favorable impression of the Watergate as compared to Milgram study (as reflected by main effects on the foreseeability of harm, the degree to which the research should have been done, and ratings of the experimenter; $ps < .05$), perhaps because of the Watergate study's direct relevance to contemporary events. These scenario differences make the absence of scenario interactions even more noteworthy and suggest high generality of the results across situations.

Relationships between Judgments

Table 4 presents correlations between judgments of how ethical and moral the experiment was and each of the other major dependent variables. These are broken down for Experiment 1 and 2, both overall and within the control condition only and for each of the four moral-position groupings obtained from the ethical positions questionnaire used in Experiment 2.

The teleological and skeptical positions suggest that the benefit/cost ratio of a study ought to be integrally tied to its ethicality; that is, the greater the amount learned from the study and the less harm done to the participants, the more moral it is. In Experiment 1, subjects' judgments of how moral and ethical the experiment was were significantly positively correlated overall with: (a) how much was learned, (b) the scientific value of the study, (c) how concerned the experimenter was about contributing something important to science, (d) how concerned the experimenter was with the dignity and welfare of the participants, (e) whether or not the research should have been done, and (f) the average evaluative rating of the experimenter. Within the control condition the correlations were generally comparable in magnitude but usually failed to reach significance due to the small cell size. However, judgments of the morality and ethicality of the experiment did not correlate significantly with cost factors such as the degree to which the experimenter could foresee harm to the participants, the amount of psychological harm that occurred, or the number of participants whom subjects predicted would have been psychologically harmed. These patterns indicate that the benefits of the experiment and the experimenter's concern for the

TABLE 4
RELATIONSHIPS BETWEEN ETHICAL JUDGMENTS AND OTHER VARIABLES

	Experiment 1		Experiment 2		Skeptics/ Pragmatists/ nonuniversal- rules (n = 46)	Teleologists/ Pragmatists/ universal- rules (n = 27)	Skeptics/ Idealists/ nonuniversal- rules (n = 46)	Deontologists/ Idealists/ universal- rules (n = 55)
	Overall correlation (n = 133)	Control condition (n = 20)	Overall correlation (n = 174)	Control condition (n = 36)				
Amount learned	.27†	.31	.28‡	.20	.24*	.56‡	.36†	.11
Scientific value	.35‡	.59†	.29‡	.16	.14	.27	.45‡	.18
Concern for science	.22*	.16	.25†	.17	.36†	.48†	.17	.15
Concern for subjects	.30‡	.14	.43‡	.15	.41†	.31*	.57‡	.29*
Should be done	.68‡	.50*	.63‡	.51‡	.52‡	.70‡	.69‡	.56‡
Rating of experimenter	.34‡	.34	.49‡	.61‡	.51‡	.35*	.59‡	.62‡
Foreseeability of harm	.03	.36	-.18*	-.03	.07	-.15	-.28*	-.22*
Threat	-.12	.00	-.36‡	-.35*	-.26*	-.18	-.54‡	-.28*
Harm	-.10	-.22	-.10	.09	.08	-.08	-.28*	.03
Predicted harm	-.07	-.09	-.22†	-.06	-.22	.08	-.32*	-.23*
Obedience	-.02	.55†	-.05	-.12	-.15	-.33*	.12	-.01
Predicted obedience	.14	.26	.06	.16	-.35†	-.05	.09	.31†

* $p < .05$.† $p < .01$.‡ $p < .001$.

participants, irrespective of harm or its foreseeability, were the primary factors that were related to the ethical judgments of these high school students.

As can be seen in the third column of Table 4, overall, the college students did not disregard costs as did the high school students. Although the benefits of the research were again positively correlated with moral judgments, cost factors such as the foreseeability of harm, the degree of threat to the dignity and welfare of the participants, and the proportion of predicted harm were negatively correlated with moral judgments. The foreseeability of harm and degree of threat correlations differ significantly ($p < .05$) from those obtained from the high school students, while the predicted harm correlations differ marginally ($p < .10$). This supports the prediction that the hedonic relevance of the scenarios brought costs more into consideration for the college than for high school students.

However, the costs were not weighted equally by all of the college subjects. The final four columns of Table 4 present correlations for subjects classified in each of the four moral-position categories described earlier. Several striking patterns are noticeable. First, pragmatists who endorse universal-rules (i.e., teleologists) exhibited a pattern of correlations that was virtually identical to that obtained from the high school students. None of the harm or threat items correlated significantly for these subjects, who apparently viewed the benefits as primary considerations. These subjects come closest to endorsing the types of beliefs held by classical teleologists; that is, costs and benefits should be weighed when making moral judgments yet some basic moral principles can be applied to most situations. The fact that their judgments did not vary with perceived harm is intriguing and supports one of Baumrind's (1971) major charges against the teleological position. Such people, Baumrind suggested, find it all too easy to justify sacrificing "the welfare of subjects in the name of science."

On the opposite side of the ledger, subjects classified as idealists who endorsed universal rules (i.e., deontologists) exhibited significant negative relationships between their moral judgments and cost factors. Yet for these subjects, none of the benefit items (amount learned, scientific value of the study, and the experimenter's concern for science) were significantly related to moral judgments. Thus, the deontologists appeared so concerned with the participants' welfare that they neglected the scientific value of the research when morally judging it. This pattern strongly supports one of Milgram's charges against Baumrind. In replying to Baumrind, Milgram (1964, p. 852) noted the existence of "a cleavage in American psychology between those whose primary concern is with *helping* people and those who are interested in *learning* about people." Our results certainly indicate that some people (deontologists) are so concerned with helping or protecting that they neglect the informational benefits of research.

The group whose moral judgments most completely covaried with both benefits and costs were the idealists who believed in nonuniversal fluid rules. These subjects are skeptics (ethical egoists) in that they take a relativistic position regarding ethical rules, yet they endorse very strict positions regarding the protection of people's welfare, apparently evidencing strict personal codes while allowing high interindividual variability. The other skeptical group, pragmatists who endorse nonuniversal rules, also balanced benefits against costs though their judgments did not covary with as many costs.

A final feature of Table 4 is the significant negative correlation obtained between predicted obedience and moral judgments for the pragmatist/nonuniversal-rules group (skeptics) and the significant positive correlation between these variables for the idealists/universal-rule group (deontologists) (the difference between these correlations is significant, $p < .001$, and each also differs significantly from those obtained in the remaining two groups, $p < .05$). The idealists/universal-rules group apparently viewed obedience as a positive behavior, perhaps one which helped the experimenter, while the pragmatists/nonuniversal-rules group apparently viewed obedience as a negative, socially undesirable behavior. This interpretation is supported by a significant negative correlation between predicted obedience and the experimenter's ability to foresee harm for the idealists/universal-rules group, $r = -.24$, $p < .05$. A nonsignificant relationship existed between these variables for the pragmatists/nonuniversal-rules group, $r = .08$, and these two correlations differed significantly. In addition, predicted obedience and predicted harm were significantly negatively correlated for both of these groups (idealists/universal-rules: $r = -.22$, $p < .05$; pragmatists/nonuniversal-rules: $r = -.26$, $p < .04$), while not for the remaining two groups (pragmatists/universal-rules: $r = -.04$; idealists/nonuniversal-rules: $r = -.02$). Thus, the idealists/universal-rules group (skeptics) viewed obedience as immoral but did not associate it with the experimenter's ability to foresee harm and even felt that it was negatively related to predictable harm; it would not be too great a departure from the data to suggest that these subjects viewed obedience negatively not because they felt it harmed the participants, but because they felt it was socially undesirable and detrimental to society.

DISCUSSION

Milgram seems to be at least partially correct in his feelings of "victimization" at the hands of his own subjects. In both Experiments 1 and 2 the frequency of obedience affected moral judgments and was directly related to the degree that the research was viewed as threatening the participants' dignity and welfare. Furthermore, ratings of the morality of

the research were positively correlated with evaluative ratings of the researcher. However, the proportion of participants who experienced psychological trauma had no significant between-conditions effects on ratings of morality and threat. The social desirability of the participants' behaviors rather than the proportion of harmed participants was critical. These effects generalized across scenarios and subject populations.

The correlational data demonstrated that the benefits and costs of the research were weighed differently by different groups when arriving at moral judgments. The high school students of Experiment 1 and the subjects classified as teleologists in Experiment 2 (i.e., pragmatist/universal-rules group) covaried their moral judgments primarily with the benefits of the research, such as its informational and scientific value, and not with its costs to participants. This pattern provides fuel for charges by the deontologists that people who espouse weighing benefits and costs actually end up neglecting the welfare of subjects when pursuing the fruits of science. The deontologists (i.e., idealists/universal-rules group), on the other hand, did the reverse, covarying their moral judgments with cost factors such as threat to participants and foreseeability of harm but not with scientific benefits. Thus it seems the deontologists, as a teleologist might suggest, protect participants at the expense of information. Also, the judgments of the deontologically oriented subjects contradict a professed assumption of their philosophy, that if a universal rule is violated, the action (experiment) is immoral, irrespective of the magnitude of the consequences. Deontologists did allow magnitudes of costs to enter into their moral calculations. Either naive deontologists differ from sophisticated philosophical adherents of that position or deontologists find it difficult to behave strictly in accord with their moral beliefs. The moral judgments of ethical skeptics (nonuniversal-rules groups) covaried more than did those of other groups with both the costs and benefits of the research; in this sense, their behavior patterns corresponded the most closely with their ideology.

These individual-difference results lead to two major implications of the present research. First, ethical philosophy can offer psychologists a great deal of insight into moral decision-making. The empirical support obtained for the major dimensions discussed by philosophers suggests the fruitfulness of greater attention to what philosophy has to offer. Second, since Heider (1958), attribution theorists who have delved into the moral judgment area have generally assumed high interpersonal consistency in people's judgments of what "ought" to be. The present results show that there are major individual differences in "ought" judgments that reflect moral philosophies and that these differences determine how people will weigh (or fail to weigh) costs and benefits when making moral judgments. These dimensions can be usefully incorporated into attribution theory to provide greater predictive power.

A major difference between the evaluative ratings of the high school and college students was on their perceptions of the informational value of the research. Obedience and psychological upset interacted to determine perceptions of informational value but did so differently for the two populations. The high school students associated the highest informational value with situations that contradicted the reward/cost structure of the conditions, feeling that more was learned when participants complied despite high costs or noncomplied despite low costs. As an attributional approach would suggest, these subjects appeared to be naively applying one of the criteria for evaluating a scientific endeavor: The more "non-obvious" the result, the more informational value it has. Judgments of the college students indicated that they thought the research was most informative when high obedience was obtained (thus making the researcher happy by justifying the research) but where few participants were upset (thus making the participants happy). The facts that high school as compared to college subjects (a) judged the informational value differently, (b) associated obedience with ethicality in the control conditions, and (c) neglected to covary participants' harm with judged ethicality suggests that hedonic relevance was more strongly involved for the college students and caused them to identify more closely with the participants. These were the only differences between the high school and college populations.

The present study represents a first step in the exploration of factors that affect ethical judgments. Given that any profession is dependent upon the good will of the general public, it is important to find out how people do evaluate psychological research and to ascertain the factors that affect such evaluations. This may or may not lead to changes in specific procedures currently used for psychological investigation. But even leaving aside the pragmatics of public and governmental support, the question of how such judgments are made is an interesting one in its own right.

REFERENCES

- American Psychological Association. *Ethical principles in the conduct of research with human participants*. Washington, D.C.: American Psychological Association, 1973.
- 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 report of the committee on ethical standards in psychological research. *American Psychologist*, 1971, **26**, 887-896.
- Blanshard, B. Morality and politics. In R. T. De George (Ed.), *Ethics and society*. Garden City, New York: Doubleday Anchor, 1966.
- Chaikin, A. L., & Darley, J. M. Victim or perpetrator? Defensive attribution of responsibility and the need for order and justice. *Journal of Personality and Social Psychology*, 1973, **25**, 268-275.

- Cook, S. W. A comment on the ethical issues involved in West, Gunn, and Chernicky's "Ubiquitous Watergate: An attributional analysis." *Journal of Personality and Social Psychology*, 1975, **32**, 66-68.
- Davis, P. E. *Introduction to moral philosophy*. Columbus, Ohio: Merrill, 1973.
- Fletcher, J. *Situation ethics*. Philadelphia: Westminster Press, 1966.
- Frankena, W. K. *Ethics*. Englewood Cliffs, New Jersey: Prentice-Hall, 1973. 2nd ed.
- Gergen, K. J. The codification of research ethics: Views of a doubting Thomas. *American Psychologist*, 1973, **28**, 907-912.
- Harvey, J. H., Harris, B., & Barnes, R. D. Actor-observer differences in the perceptions of responsibility and freedom. *Journal of Personality and Social Psychology*, 1975, **32**, 22-28.
- Heider, F. Social perception and phenomenal causality. *Psychological Review*, 1944, **51**, 358-373.
- Heider, F. *The psychology of interpersonal relations*. New York: Wiley, 1958.
- James, W. The moral philosopher and the moral life. *International Journal of Ethics*, 1891. Reprinted in P. E. Davis (Ed.), *Introduction to moral philosophy*. Columbus, Ohio: 1973.
- Jones, E. E., & Davis, K. E. From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology*. New York: Academic Press, 1965. Vol. 2.
- Kant, I. *Fundamental principles of the metaphysics of morals*. T. K. Abbott (trans.), published in 1873. Reprinted in P. E. Davis (Ed.), *Introduction to moral philosophy*. Columbus, Ohio: 1973a.
- Kant, I. *Critique of pure reason and other works on the theory of ethics*. T. K. Abbott (trans.), published in 1873. Reprinted in P. E. Davis (Ed.), *Introduction to moral philosophy*. Columbus, Ohio: 1973b.
- Kelley, H. H. Attribution theory in social psychology. *Nebraska Symposium on Motivation*, 1967, **14**, 192-241.
- Kelley, H. H. Moral evaluation. *American Psychologist*, 1971, **26**, 293-300.
- Kindler, J. M., & Gergen, K. J. An assessment of the current impact of philosophy on psychology. *American Psychologist*, 1965, **20**, 559 (Abstract).
- Kohlberg, L. The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research*. Chicago: Rand McNally, 1969.
- Milgram, S. Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, 1963, **67**, 371-378.
- Milgram, S. Issues in the study of obedience: A reply to Baumrind. *American Psychologist*, 1964, **19**, 848-852.
- Milgram, S. Some conditions of obedience and disobedience to authority. *Human Relations*, 1965, **18**, 57-76.
- Milgram, S. Interpreting obedience: Error and evidence. In A. G. Miller (Ed.), *The social psychology of psychological research*. New York: Free Press, 1972.
- Orne, M. T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 1962, **17**, 776-783.
- Phares, E. J., & Wilson, K. G. Responsibility attribution: Role of outcome severity, situational ambiguity, and internal-external control. *Journal of Personality*, 1972, **40**, 392-406.
- Plato. *The dialogues*. B. Jowett (trans.), published in 1892. Reprinted in P. E. Davis (Ed.), *Introduction to moral philosophy*. Columbus, Ohio: 1973.
- Plato. *The republic*. B. Jowett (trans.), published in 1892. Reprinted in P. E. Davis (Ed.), *Introduction to moral philosophy*. Columbus, Ohio: 1973.
- Ross, M., & Ditecco, D. An attributional analysis of moral judgments. *Journal of Social Issues*, 1975, **31**, 91-110.

- Shaver, K. G. Defensive attribution: Effects of severity and relevance on the responsibility assigned to an accident. *Journal of Personality and Social Psychology*, 1970, **14**, 101–113.
- Shaw, J. I., & Skolnick, P. Attribution of responsibility for a happy accident. *Journal of Personality and Social Psychology*, 1971, **18**, 380–383.
- Walster, E. Assignment of responsibility for an accident. *Journal of Personality and Social Psychology*, 1966, **3**, 73–79.
- West, S. G., Gunn, S. P., & Chernicky, P. Ubiquitous Watergate: An attributional analysis. *Journal of Personality and Social Psychology*, 1975, **32**, 55–65.

(Received April 21, 1976)