Rosenthal, R. & Rosnow, R.L. (2008) Essentials of Behavioral Research 3rd Ed. NY: McGraw-Hill

CHAPTER 3

ETHICAL CONSIDERATIONS, DILEMMAS, AND GUIDELINES

PUZZLES AND PROBLEMS

As mentioned in chapter 2, shortly after World War II, Karl Popper settled in England and took a position at the London School of Economics. On October 25, 1946, he gave an invited talk at Cambridge University before the Moral Science Club. Chairing the club was a renowned professor of philosophy at Cambridge, Ludwig Wittgenstein, whose views dominated British philosophy at the time (Monk, 1990). Also present was another eminent Cambridge professor of philosophy, Bertrand Russell, whose seminal work had been an early inspiration for Wittgenstein and Popper (though Wittgenstein came to regard Russell as having lost his edge and regarded his work as antediluvian). Usually at these meetings a visiting lecturer would present preliminary remarks, and Wittgenstein would then dominate the discussion. Popper and Wittgenstein harbored a deep cynicism concerning each other's views, and Russell had taken on the role of a kind of umpire at this meeting. Central heating was still virtually unknown in Britain, and the room was warmed by a coal hearth. Every so often, someone would poke the coals and clear out some of the ash in order to stir up a little more heat.

What ensued that day became the stuff of legend in philosophy. In their book entitled *Wittgenstein's Poker,* Edmonds and Eidinow (2001) recounted the controversial incident that occurred. Popper was expounding on moral philosophy when Wittgenstein, who had grabbed a red hot poker and was gesticulating with it, shouted

that Popper was confusing the issues, and challenged him to name one valid moral principle. The sequence of events is murky, but apparently Russell told Wittgenstein to put the poker down, saying to him that it was he who was confusing the issues. Popper, in response to Wittgenstein's challenge to state a valid moral principle, responded with something like, "One ought not to threaten visiting lecturers with pokers." Witnesses to the incident have never agreed on whether Wittgenstein threatened Popper (as Popper claimed), or even whether Wittgenstein was still present in the room when Popper took up the challenge. In one version of the story (which Popper repeated in his memoirs), it was Popper's pungent retort that aggravated Wittgenstein so much that he stormed out of the room, slamming the door behind him.

The controversy between Wittgenstein and Popper revolved around their different views of the proper role of philosophy. In an influential book entitled *Tractatus Logico-Philosophicus,* which first appeared in German in 1921 and was published in an English translation the following year (with an introduction by Russell), Wittgenstein had deconstructed philosophy to an atomistic level. Consisting of a series of numbered, tightly condensed, precisely articulated statements, the book begins, "The world is all that is the case," and ends with the oracular statement: "What we cannot speak about we must pass over in silence" (Wittgenstein, 1978, pp. 5, 74). Going back to the ancient Greeks, the orientation of philosophy had been the elucidation of problems—moral principles, metaphysical and epistemological issues, and so on. Wittgenstein dismissed that work as futile wordplay, contending instead that philosophers' imprecise use of ordinary language had trapped them in a bottomless pit of ambiguity. There simply are no valid *problems* in philosophy, he argued, but only linguistic *puzzles* to be resolved by revealing the misuse of language. Popper, who was as irascible as Wittgenstein, thought this argument was nonsense, and Russell had come to think that Wittgenstein's dismissal of the problem-oriented approach jeopardized the existence of philosophy as an academic discipline. Nonetheless, Wittgenstein's view dominated the British scene, and those who dared to disagree with it (like Popper) were relegated to the role of disgruntled outsiders.

Wittgenstein's and Popper's philosophical views notwithstanding, the distinction between puzzles and problems offers a convenient way to conceptualize ethical issues in science. The usual dictionary definition of a *problem* implies a dubious matter that is proposed for discussion and a solution. The problem of moral accountability in science has provoked considerable discussion and has led to a number of proposed solutions in the form of rules, regulations, and ethical guidelines. Their interpretation and implementation, however, can often be mystifying for researchers, who are obliged to *puzzle* out ways of adhering to ethical and scientific values simultaneously. We begin by giving a sense of this delicate balancing act, and throughout this chapter we mention examples of how scientists need to be attentive to societal and scientific imperatives. We also refer to the term **ethics** (derived from the Greek *ethos,* meaning "character" or "disposition"), which has to do with the values by which the conduct of individuals is morally evaluated.

Although there are a number of ethical codes in the United States and abroad, we will focus on the most prominent set of guidelines in the field of psychology, that promulgated by the American Psychological Association (APA). We review the societal context in which these guidelines were originally inspired, and we then give a flavor of the

most recent guidelines. Federal and state legal dictates imposing restrictions on the use of human subjects take precedence over the APA guidelines, but the APA ethics code (American Psychological Association, 1998) is far more focused and restrictive in many respects. Because many behavioral researchers belong not to the APA, but to the Association for Psychological Science (APS) or to other, more specialized societies (some with their own ethical guidelines, such as the Society for Research in Child Development, 1993), there is no consensus in behavioral research. Nonetheless, we will use the framework of the APA code as a way of organizing our discussion of ethical issues, including conflicts between ethical accountability and the technical demands of scientific practices. Although the primary emphasis of this discussion is on research with human subjects, we end with a brief discussion of the ethical implications of using animals in behavioral research.

A DELICATE BALANCING ACT

As Wittgenstein implied, the very language we use is loaded with traps for the unwary. For example, when clinically oriented researchers say that certain behavior is "normative" (i.e., usual or typical), the implication to the layperson is that such behavior is to be expected and is therefore desirable. When social researchers study prejudice or mental illness, they touch on highly charged societal problems. Even when researchers study topics that may seem to them to be neutral (learning behavior, for example), they must realize that to others these topics may be supercharged with values and conflicts. Thus, it seems that virtually every aspect of the research process may be viewed as value-laden to some degree, from the statement of a topic, through the conceptualization and implementation of the investigation, to the data analysis, interpretation, and reporting of findings. When research involves a societally sensitive issue (Lee, 1993), concerns about values and ethics are further heightened. To address such concerns, various national codes of ethics have been formulated by psychological associations in the United States, Canada, France, Germany, Great Britain, the Netherlands, Poland, and other countries (Kimmel, 1996; Schuler, 1982). The purpose of those codes is to provide guidelines to enable researchers to assess the morality of their scientific conduct. However, researchers must also make their way through a maze of cumbersome rules and regulations that are overseen by an independent group of evaluators, called an **institutional review board (IRB).**

As if this situation were not puzzling enough, a further problem is that ethical guidelines cannot possibly anticipate every eventual case. At best, they can provide an evolving framework for evaluating (and trying to prevent) ethical transgressions. The interpretation of the guidelines is left to IRBs, researchers, and any others who feel the need to express an opinion. Collectively, these guidelines constitute what might be described as an idealized "social contract" of do's and don'ts, to which behavioral researchers are expected to subscribe as a prerequisite of conducting any empirical studies. Broadly speaking, the agreement to which social and behavioral scientists are generally held accountable can be summed up as the responsibility (a) not to do psychological or physical harm to the subjects and (b) to do beneficent research in a way that is likely to produce valid results (Rosnow, 1997).

A further problem, however, is that, even when acting with the most noble intentions, investigators can inadvertently transgress. As philosopher John Atwell (1981)

noted, research with human subjects always "treads on thin moral ice" because investigators "are constantly in danger of violating someone's basic rights, if only the right of privacy" (p. 89). Moreover, because scientists are also ethically bound to use their abilities to advance knowledge, it has been argued that the scientific validity of the research design can also be viewed as an ethical issue, because poorly designed research cannot yield benefits and may actually be harmful (Rosenthal, 1994c, 1995a; Rosenthal & Blanck, 1993). Thus, research that is of higher scientific quality is presumed to be more ethically defensible, because of its better investment of the time of the research subjects, the funds of the granting agency, the space of the journals, and, not least, of the general investment that society has made in supporting science and its practitioners.

As we shall see in this chapter, even very experienced researchers often find themselves caught between the Scylla of scientific and theoretical requirements and the Charybdis of ethical dictates and moral sensitivities. Ironically, many seminal studies in social and behavioral science (including some of those mentioned in the previous chapter) can no longer be replicated because of obstacles imposed by daunting arrays of ethical guidelines, bureaucracies, formalities, and legalities that simply did not exist in their present form a generation or more ago (Bersoff, 1995; Fisher & Tryon, 1990; Kimmel, 1988; 1996; Koocher & Keith-Spiegel, 1990, 1998; Rosnow, Rotheram-Borus, Ceci, Blanck, & Koocher, 1993; Scott-Jones & Rosnow, 1998; Sieber, 1982a, 1982b). And yet, society and science have benefited from the accrued wisdom of those findings. A further irony is that researchers are usually held to a higher standard of accountability than are many designated and self-appointed guardians of human rights. For example, although ethical guidelines circumscribe the use of deceptive practices and the invasion of privacy, the violation of privacy as well as deception by omission (called a **passive deception**) and commission (an **active deception**) are far from rare: Lawyers routinely manipulate the truth in court on behalf of clients; prosecutors surreptitiously record private conversations; journalists often get away with using hidden cameras and other undercover practices to get stories; and police investigators use sting operations and entrapment procedures to gain the information they seek (Bok, 1978, 1984; Kimmel, 1998; Saxe, 1991; Starobin, 1997).

HISTORICAL CONTEXT OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION CODE

To put the original set of APA guidelines into context, we go back to the 1960s. During that period the American public had been whipped into a frenzy of anxiety by published reports of domestic wiretapping and other clandestine activities by the federal government. Caught up in the temper of the times, leading psychologists voiced concerns about the status of human values in research with human participants, in particular expressing disillusionment over the use of deception in social psychology (Kelman, 1967, 1968) and calling for more humanistic research methodology (Jourard, 1967, 1968). Deception was used rarely in social psychology until the 1930s, then gradually increased until the 1950s, and sharply increased in the 1950s and 1960s—and more recently there has apparently been a decline in its use (Nicks, Korn, & Mainieri, 1997). Going back to Asch's seminal studies of conformity in the 1950s, confederates had been required to deceive participants by keeping a straight face while making

ridiculous perceptual judgments. A decade later, obedience experiments done by Stanley Milgram (mentioned in Table 1.2) became the lightning rod for a heated debate about the morality of deception. Some critics argued that *any* deception in research was morally reprehensible because of its presumed adverse effects on the participants and the profession's reputation (see Kimmel, 1998, 2004, for citations and an updated discussion of this issue).

Interestingly, a subsequent survey implied that psychologists might be more concerned about ethical issues in research than were their typical participants (Sullivan & Deiker, 1973). Adding fuel to the debate were more shocking events elsewhere. In biomedical research, flagrant abuses—some resulting in the death of the human participants—were uncovered (Beecher, 1966, 1970). A notorious case, not made public until 1972, involved a U.S. Public Health Service study, conducted from 1932 to 1972, of the course of syphilis in more than 400 low-income African American men in Tuskegee, Alabama (Jones, 1993). The participants, who had been recruited from churches and clinics, were not told they had syphilis but were only told they had "bad blood." Nor were they given penicillin when it was discovered in 1943. They were given free health care and a free annual medical examination but were told they would be dropped from the study if they sought treatment elsewhere. The Public Health Service officials went so far as to have local physicians promise not to give antibiotics to subjects in the study (Stryker, 1997). As the disease progressed in its predictable course without treatment, the subjects experienced damage to the skeletal, cardiovascular, and central nervous systems and, in some cases, death. The Tuskegee study was not halted until 1972, when details were made public by a lawyer who had once been an epidemiologist for the Public Health Service. Among the horrendous abuses in this study were that subjects were not informed of the nature of the inquiry or the fact that their disease was treatable by medical care readily available at that time (Fairchild & Bayer, 1999).

Already in the 1960s, however, there were demands for reforms, with issues of research abuses and misconduct raised in newspapers, magazines, and congressional hearings (Kelman, 1968). For some time the APA (in its code of professional ethics) had addressed issues such as the confidentiality of research data. Spurred on by eloquent spokespersons who called for the codification of the research methods used in psychological research (e.g., M. B. Smith, 1967, 1969), the APA in 1966 created a task force—called the Cook Commission, after Stuart W. Cook, its chair—that was assigned to write a code of ethics for research with human participants. Out of those deliberations came a 1971 draft report (Cook et al., 1971) and a revised report in 1972 (Cook et al., 1972). The complete code was formally adopted by the APA in 1972, reissued a decade later (American Psychological Association, 1982), and in the late 1990s rewritten by a task force, which for a time was cosponsored by the APA and the APS. After a disagreement about the spirit and content of the draft report (American Psychological Association, 1998), the APS withdrew its collaboration; a draft report was then circulated by the APA alone (American Psychological Association, 1998).

Table 3.1 lists the ten ethical guidelines representing the core of the requirements that appeared in the 1982 version of the APA code. Drawing from philosophy, law, and the American experience, European psychologists had by the early 1980s formulated their own codes of ethical principles to help them meet their responsibilities to subjects (Schuler, 1981). Three principles that appeared without

TABLE 3.1

Ethical principles for research with human participants

The decision to undertake research rests on a considered judgment by the individual psychologist about how best to contribute to psychological science and human welfare. Having made the decision to conduct research, the psychologist considers alternative directions in which research energies and resources might be invested. On the basis of this consideration, the psychologist carries out the investigation with respect and concern for the dignity and welfare of the people who participate and with cognizance of federal and state regulations and professional standards governing the conduct of research with human participants.

A. In planning a study, the investigator has the responsibility to make a careful evaluation of its ethical responsibility. To the extent that the weighing of scientific and human values suggests a compromise of any principle, the investigator incurs a correspondingly serious obligation to seek ethical advice and to observe stringent safeguards to protect the rights of human participants.

B. Considering whether a participant in a planned study will be a "subject at risk" or a "subject at minimal risk," according to recognized standards, is of primary ethical concern to the investigator.

C. The investigator always retains the responsibility for ensuring ethical practice in research. The investigator is also responsible for the ethical treatment of research participants by collaborators, assistants, students, and employees, all of whom, however, incur similar obligations.

D. Except in minimal-risk research, the investigator establishes a clear and fair agreement with research participants, prior to their participation, that clarifies the obligations and responsibilities of each. The investigator has the obligation to honor all promises and commitments included in that agreement. The investigator informs the participants of all aspects of the research that might reasonably be expected to influence willingness to participate and explains all other aspects of the research about which the participants inquire. Failure to make full disclosure prior to obtaining informed consent requires additional safeguards to protect the welfare and dignity of the research participants. Research with children or with participants who have impairments that would limit understanding and/or communication requires special safeguarding procedures.

E. Methodological requirements of a study may make the use of concealment or deception necessary. Before conducting such a study, the investigator has a special responsibility to (1) determine whether the use of such techniques is justified by the study's prospective scientific, educational, or applied value; (2) determine whether alternative procedures are available that do not use concealment or deception; and (3) ensure that the participants are provided with sufficient explanation as soon as possible.

F. The investigator respects the individual's freedom to decline to participate in or to withdraw from the research at any time. The obligation to protect this freedom requires careful thought and consideration when the investigator is in a position of authority or influence over the participant. Such positions of authority include, but are not limited to, situations in which research participation is required as part of employment or in which the participant is a student, client, or employee of the investigator.

G. The investigator protects the participant from physical and mental discomfort, harm, and danger that may arise from research procedures. If risks of such consequences exist, the investigator informs the participant of that fact. Research procedures likely to cause serious or lasting harm to a participant are not used unless the failure to use these procedures might expose the participant to risk of greater harm or unless the research has great potential benefit and fully informed and voluntary consent is obtained from each participant. The participant should be informed of procedures for contacting the investigator within a reasonable time period following participation should stress, potential harm, or related questions or concerns arise.

H. After the data are collected, the investigator provides the participant with information about the nature of the study and attempts to remove any misconceptions that may have arisen. Where scientific or humane values justify delaying or withholding this information, the investigator incurs a special responsibility to monitor the research and to ensure that there are no damaging consequences for the participant.

I. Where research procedures result in undesirable consequences for the individual participant, the investigator has the responsibility to detect and remove or correct these consequences, including long-term effects.

J. Information obtained about a research participant during the course of an investigation is confidential unless otherwise agreed upon in advance. When the possibility exists that others may obtain access to such information, this possibility, together with the plans for protecting confidentiality, is explained to the participant as part of the procedure for obtaining informed consent.

Note: From *Ethical Principles in the Conduct of Research with Human Participants,* 1982, Washington, DC, pp. 5–7. Used by permission of the American Psychological Association.

exception in all of the European and American codes were (a) to avoid physical harm, (b) to avoid psychological harm, and (c) to keep the data confidential (Schuler, 1982). The third principle, which evolved to safeguard the information divulged by clients in clinical situations, was commonly justified on the basis of three claims: (a) that fairness required respect for the research participants' privacy, (b) that scientists had the professional right to keep such disclosures secret, and (c) that more honest responding by subjects should result when the investigator promised to keep the subjects' personal disclosures confidential (Blanck, Bellack, Rosnow, Rotheram-Borus, & Schooler, 1992; Bok, 1978). Despite the value of those guidelines, professional codes had not incorporated much in the way of penalties for noncompliance. The negative sanction for violating the APA ethical code was censure or expulsion from the APA—by no means considered a severe penalty, as many psychologists engaged in productive, rewarding research careers do not belong to the APA.

The 10 guidelines in Table 3.1 were formulated with the aim of instructing psychological researchers about what their moral responsibilities are, how to decide what aspects of a proposed study might pose ethical risks, and how to choose an ethical strategy for addressing such problems. Notice, for example, that Principle E does not prohibit deception; instead, it implies when a deception may be permissible and also the attendant ethical responsibilities of researchers who want to use a deception. In fact, by the time of the first adoption of the APA research code, an assortment of deceptions had slipped into many researchers' methodological arsenals (Arellano-Galdames, 1972; Gross & Fleming, 1982). Active deceptions included misrepresenting the purpose of the study or the identity of the investigators, falsely promising something to subjects, misrepresenting the equipment or procedures, and using placebos, pseudosubjects, and secret treatments. Passive deceptions included disguising experiments in natural settings, observing people in a public setting without telling them they were being studied, secretly recording potentially embarrassing behavior, and using projective tests and other instruments without disclosing their purpose to the participants.

THE BELMONT REPORT, FEDERAL REGULATIONS, AND THE INSTITUTIONAL REVIEW BOARD

A moment ago we alluded to a survey that showed psychologists to be more concerned about ethical sensitivities than were their typical participants (Sullivan & Deiker, 1973). Not every person felt the urgent need to codify such sensitivities, however. For example, Kenneth Gergen (1973a) expressed another popular sentiment among researchers when he warned of the possibility of a precarious trade-off of scientific advances for excessive constraints:

> Most of us have encountered studies that arouse moral indignation. We do not wish to see such research carried out in the profession. However, the important question is whether the principles we establish to prevent these few experiments from being conducted may not obviate the vast majority of contemporary research. We may be mounting a very dangerous cannon to shoot a mouse. (p. 908)

A few years later, however, what Gergen characterized as a "dangerous cannon" seemed more like a popgun in light of dramatic changes that occurred when the review process was set in motion in 1974 by the National Research Act (Pub. L. No. 93-348). That statute also created the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research.

The National Commission conducted hearings over a 3-year period, which culminated in the **Belmont Report,** named from the discussions that were held at the Smithsonian Institution's Belmont Conference Center in Washington, DC (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979). Ethical guidelines, the report concluded, should emphasize (a) showing respect for individuals as autonomous agents and the protection of those with diminished autonomy, (b) maximizing plausible benefits and minimizing possible harms, and (c) using fairness or justice in distributing risks and benefits. In addition, federal directives now ordered institutions applying for grant support to create review boards to evaluate grant submissions (e.g., Department of Health, Education, and Welfare, 1978). If participation in a research study is classified by the IRB as involving more than "minimal risk," that study requires the use of specific safeguards. The safeguards include providing the participants with an adequate explanation of the purposes of the research, the procedures to be used, the potential discomforts and risks to subjects, the benefits that subjects or others may receive, the extent of anonymity in any records that are kept, and the identity of an individual that subjects can contact about the research (Delgado & Leskovac, 1986). Most important perhaps is the investigator's responsibility to make sure that participants understand their prerogative to withdraw from the study at any time without penalty. The spirit of the federal dictates is the same as that of the APA guidelines (Table 3.1), except that the government's rules are legally enforceable in a significant way.

Only a few years after they were created, IRBs had become a source of consternation to many researchers, who felt their research "had been impeded in a way that was not balanced by the benefits of the review process" (Gray & Cook, 1980, p. 40). In recent years, particularly with the development of research on AIDS (acquired immune deficiency syndrome), the sphere of responsibility of IRBs has been expanded as a result of a proliferation of self-imposed safeguards, legally mandated constraints, pressures by advocacy groups, and methodological innovations. The responsibility of IRBs is no longer limited to the evaluation of grant submissions or funded research and may encompass any proposed study in an institution. **Minimal risk research** (i.e., studies in which the likelihood and extent of harm to the subjects is perceived to be no greater than that typically experienced in everyday life or in routine physical or psychological examinations or tests) is authorized to get an expedited review, but even the most innocent study can touch a nerve in some designated regulatory body. Not many years ago, IRBs were seen as the guardians of informed consent, confidentiality, and the safety and autonomy of the research participants. Today, some IRBs, particularly in medical schools, evaluate technical and statistical aspects of research.

As if the pursuit of behavioral research were not already complicated, there are also state laws that limit the type of information requested of participants and the degree of acceptable risk to them, implying that some IRBs are legally bound to impose stricter standards. It is not uncommon that a research proposal approved without alterations at

one institution will be substantially modified or even rejected by an IRB at another participating institution (Ceci, Peters, & Plotkin, 1985; P. C. Williams, 1984). The problem of variability in decision making about research on sensitive issues is compounded by the subjectivity of an ethical review and the individual biases of IRB members (Kimmel, 1991). As Ceci et al. (1985) noted, getting a socially sensitive proposal approved is sometimes a matter of the luck of drawing a particular group of IRB members whose values just happen to be congruent with the values of the researchers. In light of these developments, the latest version of the APA research code is self-described as more "aspirational" than "prescriptive," although there are, of course, certain behaviors that cannot be condoned under any circumstances (e.g., fraud). The emphasis of the current APA code is on five broad principles (reflecting the spirit of both the Belmont Report and various federal statutes and directives), which we explore next: (a) respect for persons and their autonomy, (b) beneficence and nonmaleficence, (c) justice, (d) trust, and (e) fidelity and scientific integrity (American Psychological Association, 1998).

PRINCIPLE I: RESPECT FOR PERSONS AND THEIR AUTONOMY

The first principle of the current APA document is a reflection of the earlier Principle D (see Table 3.1) and implies that our ethical (and legal) responsibility is to ensure that people's privacy is adequately protected, that potential participants know what they will be getting into, that they will not be humiliated, and that they are free to decide whether or not to participate. The heart of this principle is **informed consent,** which refers to the procedure in which prospective subjects (or their legally authorized representatives or guardians) voluntarily agree to participate in the research after being told about its purpose, including the nature of the instruments to be used and any anticipated risks and benefits (Scott-Jones & Rosnow, 1998). To the extent they are capable, prospective participants must be given the opportunity to choose what shall or shall not happen to them. If they have diminished autonomy (e.g., because of immaturity, incapacitation, or other circumstances that limit or restrict their ability or opportunity for autonomous choice), or if they have difficulty understanding the nature of the research because they are young or feeling anxious (Dorn, Susman, & Fletcher, 1995; Susman, Dorn, & Fletcher, 1992), then they must be appropriately represented and protected. The responsibility for obtaining legally effective informed consent is the obligation of the principal investigator (Delgado & Leskovac, 1986).

For example, whenever children or adolescents are proposed as subjects, researchers are required to obtain legally effective parental consent before proceeding and are not permitted to make appeals to children to participate before parental consent is obtained. If the children do not live with their parents (e.g., are wards of some agency), the researcher can speak with an advocate who is appointed to act in the best interests of the child in the consent process. Once informed consent of the parent or advocate has been obtained, the researcher asks the child on the day of the study whether he or she wishes to participate—assuming the child is mature enough to be asked about participation (Scott-Jones & Rosnow, 1998). It has been noted, however, that an unfortunate consequence of increased scrutiny by IRBs is that the disclosure procedure has become so detailed and cumbersome in many institutions that it may actually defeat

the purpose for which informed consent was intended (Imber et al., 1986). One psychologist reported that many of his adult subjects mistakenly thought they had relinquished their legal protection by signing an informed consent agreement (Mann, 1994), although the right to sue for negligence is protected by federal regulations on the use of human participants (Department of Health and Human Services, 1983).

A further concern is that, sometimes, informing subjects of some pertinent aspect of the investigation may impair the validity of the research. For example, Gerald T. Gardner (1978) performed a series of studies of the effects of noise on task performance. The aim of the research was to replicate a phenomenon first reported by Glass and Singer (1972), indicating that exposure to uncontrollable, unpredictable noise can negatively affect task performance. Although Gardner's initial experiments duplicated Glass and Singer's findings, two subsequent experiments did not. Bewildered by that outcome, Gardner sought to puzzle out a reason for the discrepancy. The only difference in procedure between the early and later studies in the series was that the first studies had been performed before the implementation of federal guidelines requiring informed consent, and the later studies had been carried out using informed consent. This difference inspired Gardner to hypothesize that informed consent might actually have been responsible for the discrepant results.

Acting on this hypothesis, Gardner conducted a final study in which two groups were exposed to uncontrollable noise; one group had given informed consent, whereas the other group had not. The results of this study were that the group that had given informed consent did not show the emergence of negative effects of the noise, but the other group did. Gardner reasoned that negative effects did not emerge because the informed consent had created a perception in the subjects of control over the noise. As Gardner (1978) explained, perceived control "could result from references . . . in the consent form to subjects' ability to withdraw from the study without penalty, to their freedom to choose an alternative to [subject] pool participation" (p. 633). Apparently, conforming to the new ethical guidelines in this instance seriously impaired the emergence of the negative effects of laboratory stressors. Had federal guidelines been instituted when Glass and Singer initiated their research in the late 1960s, is it possible that important facts about environmental noise would never have come to light?

Another early study was performed by clinical researchers Jerome H. Resnick and Thomas Schwartz (1973), who suspected that, in some circumstances, informed consent might trigger "paranoid ideation in otherwise nonsuspicious subjects" (p. 137). Using the traditional verbal conditioning procedure described in chapter 1 (used by Crowne and Marlowe), Resnick and Schwartz experimentally manipulated the ethical standard of informed consent. The subjects were presented with a sequence of cards, each of which showed a specific verb and six pronouns *(I, you, we, he, she, they)* and were told to make up a sentence using the verb and any of the six pronouns. They were then verbally reinforced by the experimenter, who said "good" or "ok" each time the subject chose either *I* or *we*. Before the study began, half the prospective subjects were told the nature of the conditioning procedure in strict adherence with informed consent guidelines; the control subjects were not given that information, but were run just as the study would have been conducted before the era of informed consent.

The principal finding was that the control group conditioned as expected, but the fully informed subjects exhibited an unexpected reversal in the pattern of their conditioning behavior. Using postexperimental questionnaires, Resnick and Schwartz discovered that many of the fully informed subjects, after having been told so much about the study, questioned in their own minds the experimenter's "true" hypothesis. One subject stated that he "had wanted to play it cool; and to give the impression that the experimenter's reinforcements were having no effect" (p. 138). When told that his use of the two reinforced pronouns had decreased by more than half from the first 20 trials to the last 20, this person laughed and said, "I was afraid I would overdo it" (p. 138). Not only was it distressing to Resnick and Schwartz that their fully informed subjects were distrustful, but it was unclear what was happening in, as these researchers put it, "a room full of mirrors where objective reality and its perception blend, and thereby become metaphysical" (p. 138). The results seemed to imply that standard textbook principles of verbal learning would turn backward if all previous studies in this area had strictly adhered to fully informed consent. This study raised a red flag signaling that full disclosure may sometimes be an impediment to the pursuit of knowledge.

Thus, we see that, as Gergen (1973a) and others anticipated, there are scientific puzzles associated with strict compliance with informed consent. In a later chapter we allude to another potential concern, which is the "delicate balance" between *experimenter and subject artifacts* (i.e., specific threats to validity that can be attributed to uncontrolled researcher- or participant-related variables) and ethics in behavioral research (Rosnow & Rosenthal, 1997; Suls & Rosnow, 1981). In chapter 9 we will describe how using volunteer subjects could introduce biases that then make the research results more difficult to generalize to populations consisting in part of potential nonvolunteers (Rosenthal & Rosnow, 1969b, 1975; Rosnow & Rosenthal, 1970, 1976).

PRINCIPLE II: BENEFICENCE AND NONMALEFICENCE

Beneficence means the "doing of good," which implies that the research is expected to have some conceivable benefit, and **nonmaleficence** implies that, as in the Hippocratic oath that physicians take, behavioral and social researchers are also expected to "do no harm." The avoidance of harm as a standard for ethical research originally emanated from the Nuremberg Code of 1946–1949, developed in conjunction with expert testimony against Nazi physicians at the Nuremberg Military Tribunal after World War II. The risks of behavioral and social research pale by comparison with the appalling "experiments" done by Nazi physicians in the name of science, but federal regulations nevertheless insist that assessment of risk be part of the ethical evaluation of all proposed research with human subjects. Generally speaking, the most significant risks in traditional psychological research are associated with privacy invasion or the use of some active or passive deception. When deception is used, the assumption is that (a) the research has genuine scientific value, (b) providing the subjects with full details of the research would seriously impair its validity, (c) no undisclosed "risks" to the subjects are more than minimal, and (d) the subjects will be adequately debriefed at some appropriate time.

Prior to the Belmont Report, a classic example of behavioral research that became the focus of concerns about the use of deception was social psychologist Stanley Milgram's research on how far a person would go in subjecting another person to pain by the order of an authority figure (Milgram, 1963, 1965). Volunteer subjects, placed in the role of the "teacher," were deceived into believing that they would be giving varying degrees of painful electric shocks to another person (the "learner") each time he made a mistake in a learning task. Milgram varied the physical proximity between the teacher and the learner, to see whether the teacher would be less ruthless in administering the electric shocks as he or she got closer to the learner. The results were that a great many subjects (the teachers) unhesitatingly obeyed the researcher's command as they continued to increase the level of shock administered to the learner. Even when there was feedback from the learner, who pretended to cry out in pain, many subjects obeyed the researcher's order to "please continue" or "you have no choice, you must go on." The subjects were not told at the outset that the shock apparatus was fake but were extensively debriefed once the experiment was over. Even though the learner was a confederate of Milgram's and there were no actual shocks transmitted, concerns about ethics and values have dogged these studies since they were first reported (Milgram, 1963, 1965, 1975, 1977).

For instance, psychologist Diana Baumrind (1964) quoted Milgram's own descriptions of the reactions of some of the subjects:

> I observed a mature and initially poised businessman enter the laboratory smiling and confident. Within 20 minutes he was reduced to a twitching, stuttering wreck, who was rapidly approaching a point of nervous collapse. He constantly pulled on his earlobe, and twisted his hands. At one point he pushed his fist into his forehead and muttered: "Oh God, let's stop it." And yet he continued to respond to every word of the experimenter and obeyed to the end. (Milgram, 1963, p. 377)

Baumrind posed the question of why Milgram had not terminated the deception when he saw that it was so stressful to his subjects. She concluded that there could be no rational basis for doing this kind of research, unless the subjects were forewarned of the psychological risks. Another criticism was that Milgram's deception had instilled in his subjects a general distrust of authority, and thus the study was unethical no matter whether the subjects were debriefed afterward.

Milgram (1964) responded that it was not his intention to create stress, and, further, that the extreme tension induced in some subjects had not been expected. He noted that, before carrying out the research, he had asked professional colleagues for their opinions, and none of the experts anticipated the behavior that subsequently resulted. He stated that he also thought the subjects would refuse to follow orders. In spite of the dramatic appearance of stress, he believed there were no injurious effects to the subjects. Each subject was shown that the learner had not received dangerous electric shocks but had only pretended to receive them. Milgram also sent questionnaires to the subjects to elicit their reactions after they had been given a full report of his investigation. Less than 1 percent said they regretted having participated, 15 percent were neutral or ambivalent, and over 80 percent responded that they were glad to have participated. As for the criticism that his use of deception

had instilled a general distrust of authority, he replied that the experimenter in his research was not just *any* authority, but someone who ordered the subjects to act harshly and inhumanely toward another person. Milgram added that he would consider the result of the highest value if participation in the research did indeed inculcate a skepticism of that kind of authority.

Duping subjects into believing they were administering painful electric shocks to another person is inherently disquieting, but it is hard to imagine how forewarning them about the use of deception would not have destroyed the validity of the investigation. Furthermore, Milgram's follow-up treatments were unusually extensive. During the postexperimental debriefing session, he made sure that each subject was shown the reality of the experimental situation and had a friendly reconciliation with the learner and an extended discussion with the experimenter about the purpose of the study and why it was thought necessary to use deception. Subjects who had obeyed the experimenter when ordered to keep administering the electric shocks were told that their behavior was not abnormal, and that the feelings of conflict or tension they had experienced were shared by other subjects. The subjects were told that they would receive a comprehensive written report at the conclusion of the study. The report they received detailed the experimental procedures and findings, and the subject's own part in the research was treated with dignity. Subjects also received a questionnaire that asked them once again to express their thoughts and feelings about their behavior. One year after the experiment was completed there was an additional follow-up of 40 of the experimental subjects, who were intensively interviewed by a psychiatrist in order to rule out any delayed injurious effects resulting from the experiment.

Milgram's follow-up treatments were more comprehensive than is characteristic of most studies. Subsequent ethical guidelines call for **debriefing** if deception is used in research (also referred to as *dehoaxing*) in order to remove any misconceptions the subjects may have about the research, allay any negative emotions or thoughts, and leave them with a sense of dignity, knowledge, and a perception of time not wasted. That debriefing can also provide researchers with information that subjects may be either reluctant or unable to disclose at any other point in the study (Rotheram-Borus, Koopman, & Bradley, 1989). For example, in experimental trials with persons infected by HIV, it has been a common practice for many participants to share medication with each other, gain access to drugs or treatments available outside the study, and take multiple drugs simultaneously, thereby making it almost impossible to conduct an evaluation of a single drug uncontaminated by auxiliary treatments (Blanck et al., 1992). Debriefing in this situation includes monitoring the degree and type of multiple drug use among subjects in the trials. Jones and Gerard (1967) suggested that debriefing also include discovery about what each subject thought of the research situation, providing the investigator with an experiential context in which to interpret the results.

As mentioned, if the research involves any sort of deception, debriefing is usually expected to be used to reveal the truth about the study and the careful consideration that has been given to the use of the deception. For example, it might be explained to subjects that science is the search for truth, and that sometimes it is necessary to resort to withholding information in order to uncover the truth. In some cases, however, the revelation that a deception was part of the study spawns skepticism and leaves the subjects feeling gullible, as if they have been "had" by a fraudulent

procedure. Thus, it is also important to weigh the welfare and rights of the participants against the possibility that dehoaxing might itself lead to psychological discomfort (Fisher & Fyrberg, 1994). Assuming that debriefing can be done, the researcher might explain that being "taken in" does not reflect in any way on the subject's intelligence or character but simply shows the effectiveness or validity of the research design. Presumably, the researcher took some pains to achieve an effective design so as not to waste the subjects' time and effort. Most important, the debriefing should proceed gradually and patiently, with the chief aim of gently unfolding the details of any deceptions used and reducing any negative feelings. Instead of thinking of themselves as "victims," the subjects should then more correctly realize that they are "coinvestigators" in the search for truth (Aronson & Carlsmith, 1968; Mills, 1976; Rosnow & Rosenthal, 1997).

PRINCIPLE III: JUSTICE

The third principle, simply called **justice,** implies "fairness" and, in behavioral research, refers to the ideal that the burdens as well as the benefits of the scientific investigation should be distributed equitably. The men who participated in the Tuskegee study could not have benefited in any significant way, and they alone bore the awful burdens as well. However, suppose it had been an experiment to test the effectiveness of a new drug in curing syphilis, the strategy being to give half the men at random the new drug and the other half a fake "pill" masquerading as the real thing (i.e., a **placebo**). Would that approach have made the study any more acceptable? Research on AIDS has made investigators sensitive to such ethical questions, and one response is to include potential participants or surrogates for them in the decision-making process—although this inclusion does not absolve investigators themselves of their own responsibilities to protect the safety and rights of their subjects (Melton, Levine, Koocher, Rosenthal, & Thompson, 1988). If an effective treatment is available, use of the effective treatment can be the control condition, so that the experimental comparison is now between the new therapy and the effective alternative. The Declaration of Helsinki, adopted by the general assembly of the World Medical Association in 2000, stipulated that a placebo be used only when there is no other effective drug or therapy available for comparison with the therapeutic being tested. In the case of the Tuskegee study, there was an effective treatment available (penicillin), and depriving men of that treatment made the study profoundly unjust. Another design alternative (discussed in chapter 7), which is useful in certain randomized experiments (the Tuskegee study was not a randomized experiment), is to use a **wait-list control group;** in such a design the alternative therapy is given to the control group after it has been administered in the experimental group and the results have been documented.

As daily life constantly reminds us, however, social, political, and legal justice are ideals that are unlikely to be achieved in a world that is never fully just. Is fairness or justice, then, merely in the eyes of the beholder? Philosophers make a distinction between two orientations, the consequentialist and the deontological, and argue that how people view ethical questions depends on their orientation. The **consequentialist view** refers to the argument that whether an action is right or wrong depends on its consequences. The **deontological view** is that some actions may be presumed to be

categorically wrong no matter what their consequences (e.g., threatening a visiting lecturer with a red hot poker). In fact, there is empirical support for the idea that people tend to judge the world from one of these two perspectives, or from a pluralistic orientation that encompasses aspects of both (Forsyth, 1980; Forsyth & Pope, 1984). Milgram lied to his subjects, and that lying was immoral if we believe that lying in *any* form is wrong (a deontological argument). On the other hand, it would appear that Baumrind's views of Milgram's research were influenced by her awareness of his results (the consequentialist view), just as Milgram's ideas may have been colored by his own pluralistic approach (i.e., containing elements of both the consequentialist and the deontological views, but not a blanket condemnation of deception). But some have also argued that deception was not all that was at stake. The studies were "unjust" because Milgram exposed his subjects to a possibility of unwanted and unasked-for self-knowledge (Cassell, 1982). How we ourselves perceive those issues may be a window into the nature of our personal orientation as consequentialist, deontological, or pluralistic (Forsyth & Pope, 1984; C. P. Smith, 1983; Waterman, 1988).

Another early study helped to underscore the problem that injustice is not always easily anticipated. It involved a 1973 field experiment designed in part to improve the quality of work life at the Rushton Mining Company in Pennsylvania (Blumberg, 1980; Susman, 1976). Developed on the basis of previous research in the United Kingdom (Trist & Bamforth, 1951; Trist, Higgin, Murray, & Pollock, 1963), the Rushton project had as its specific aims to improve employee skills, safety, and job satisfaction while raising the level of performance and company earnings (Blumberg & Pringle, 1983). After months of preparation by the researchers and the mining company, a call was issued for volunteers for a work group that would have direct responsibility for the production in one section of the mining operations. The volunteers were instructed to abandon their traditional roles and, after extensive training in safety laws, good mining practices, and job safety analysis, were left to coordinate their own activities. Paid at the top rate, that of the highest skilled job classification in that section, they became enthusiastic proponents of "our way of working."

All was not so rosy in the rest of the mine, however. Other workers, those in the control condition, expressed resentment and anger at the "haughtiness" of the volunteers and the injustice of the reward system. The volunteers had even been treated to a steak and lobster dinner by the president of the company, the others complained. Why should these "inexperienced" workers receive special treatment and higher pay than other miners with many more years on the job? Rumors circulated through the mine that the volunteers were "riding the gravy train" and being "spoon-fed," and that autonomy was a "communist plot" because all the volunteers received the same rate and the company was "making out" at their expense. The researchers were rumored to be politically motivated to "bust the union" (Blumberg & Pringle, 1983). No matter what the important theoretical and applied benefits of the research would have been, the seeds of conflict were planted, and the experiment had to be prematurely concluded.

In this case we see that applied research can have its own problems and puzzles, quite apart from those encountered by Milgram. There was no deception or invasion of privacy in the Rushton study, but there was the problem of "injustice" because a sizable number of workers (nonvolunteers, to be sure) did not receive the benefits enjoyed by those in the experimental group. Still other risks may occur in applied

research. For example, a moral cost may be involved simply in the publication of the results. That is, the report might (a) upset some persons who are able to identify themselves in the publication, (b) subject the community to possible embarrassment or to unwanted publicity, (c) make those who are identifiable vulnerable to others who have power over them, or (d) weaken the scientific enterprise by communicating to people that science is exploitive (Johnson, 1982). On the other hand, what would be the social and scientific costs of *not* disseminating research findings? In chapter 1 we listed the ability to communicate—and by extension, the written record—as one of the essentials of sound scientific practice. The written record of the search for truth is the official archive that tells us about the observations that were made, the hypotheses that were examined (and those that were ignored), the ideas that were found wanting, and those that withstood the test of further observations. One author was quoted in chapter 1: "Scientists must write . . . so that their discoveries may be known to others" (Barrass, 1978, p. 25).

Furthermore, "unfairness" and "injustice" are hardly limited to research situations. For example, Broome (1984) discussed the ethical issue of fairness in selecting people for chronic hemodialysis—a medical procedure that can save the life of a person whose kidneys have failed. It is expensive, and in many countries there are just not enough facilities available to treat everyone who could benefit. Because without treatment a patient quickly dies, how should a candidate for hemodialysis be selected? First come, first served was one way that some hospitals chose candidates. The inventor of hemodialysis, B. H. Scribner, is said to have selected people on the basis of their being under 40 years old, free of cardiovascular disease, pillars of the community, and contributors to the community's economics. He is also said to have taken into account whether they were married and whether they went to church. Still another approach uses randomness. Broome (1984) pointed out that selecting people randomly—such as by using a lottery to choose conscripts to fight in a war—is often justified as the "fairest" procedure because everyone has an equal shot at being selected for life or death. But suppose conscripts for the military were instead selected not randomly, but on the grounds of who was the biggest and strongest? Which approach is fairer: randomness or selection on the grounds of who is more likely to survive? Some hospitals chose candidates for hemodialysis on the basis of a lottery among those patients who were judged to be most medically suitable.

PRINCIPLE IV: TRUST

This principle refers to the establishment of a relationship of trust with the participants in the study. It is based on the assumption that subjects are fully informed about what they will be getting into, that nothing is done to jeopardize this trust, and that their disclosures are protected against unwarranted access. This last requirement is what is meant by **confidentiality,** which is intended to ensure the subjects' privacy by setting in place procedures for protecting the data. For example, the investigator might use a coding system in which the names of the participants are represented by a sequence of numbers that are impossible for anyone else to identify. In cases in which participants respond anonymously and are never asked to give any information that would identify them, their privacy is obviously protected. In certain government-funded

biomedical and behavioral research, it is sometimes possible for researchers to obtain from the funding agency a **certificate of confidentiality,** which is a formal agreement that requires the researcher to keep the data confidential and thus exempts the data from subpoena. However, the extent to which such a certificate can actually provide legal protection has not yet been established in the courts, and it is further complicated by the existence of laws that require the reporting of certain sensitive information.

For example, as stipulated by the Child Abuse Prevention and Treatment and Adoption Reform Act (1992) and its revisions and amendments (Pub. L. No. 102-295), each state must pass laws to require the reporting of child abuse and neglect. The nature and wording of the statutes is left to the discretion of the states, which have increasingly expanded the lists of individuals who are obligated to report suspected cases (Liss, 1994). As a result, developmental psychologists working in the area of intervention research are often pressed to report child abuse. Researchers who are investigating abuse may be torn between reporting a suspected culprit (jeopardizing the relationship of trust?) and losing a valuable participant in the study (jeopardizing the validity or generalizability of the study?), or they may feel they do not have the moral right to report a parent on the basis of their limited training and the evidence they have. It may be that charges of abuse will not be proven, but this possibility does not excuse researchers from their legal responsibilities (Liss, 1994). Researchers who lack the training and clinical acumen to recognize abuse may overreport suspected cases (Scott-Jones, 1994). These are obviously knotty problems, and they have led to suggestions about the need for specialized training opportunities for some researchers, new reporting methods in research protocols, the restructuring of ethical guidelines in populations at risk, and further research on the predictive power and diagnostic validity of the relevant assessment tools (e.g., C. B. Fisher, 1994; Scarr, 1994; Scott-Jones, 1994).

PRINCIPLE V: FIDELITY AND SCIENTIFIC INTEGRITY

The relationship between scientific quality and ethical quality is the essence of the fifth principle, which includes a wide range of issues (Rosenthal, 1994c, 1995a; Rosenthal & Blanck, 1993). One issue involves telling prospective participants, granting agencies, colleagues, administrators, and ourselves that the research is likely to achieve goals that it is, in fact, unlikely to achieve, that is, *hyperclaiming* (Rosenthal, 1994c). It is true that colleagues can often figure out for themselves whether research claims are exaggerated, but prospective subjects are not usually equipped to question "hyperclaims," such as the idea that an investigation will yield a cure for panic disorder, depression, schizophrenia, or cancer. Closely related to this problem is **causism,** which means implying a causal relationship where none has been established by the available data. Characteristics of this problem include (a) the absence of an appropriate evidential base; (b) the use of language implying cause (e.g., "the effect of," "the impact of," "the consequence of," "as a result of") where the appropriate language would actually be "was related to," "was predictable from," or "could be inferred from"; and (c) self-serving benefits, because it makes the causist's result appear more important or fundamental than it really is (Rosenthal, 1994c). A perpetrator of causism who is unaware

of the hyperclaim shows poor scientific training or lazy writing. The perpetrator who is aware of the hyperclaim shows blatant unethical misrepresentation and deception.

To illustrate how a poorly trained investigator might stumble into causism, imagine that a research protocol that comes before an IRB proposes the hypothesis that private schools improve children's intellectual functioning more than public schools do. Children from randomly selected private and public schools are to be tested extensively, and the research hypothesis is to be tested by a comparison of the scores earned by students from private and public schools. The safety of the children to be tested is certainly not an issue; yet it can be argued that this research violates the principle of scientific integrity because of the inadequacy of the research design. The purported goal of the study is to learn about the "causal impact on performance of private versus public schooling," but the design of the research does not permit sound causal inference because of the absence of random assignment to conditions or of a reasonable attempt to consider plausible rival hypotheses (Cook & Campbell, 1979). The design provokes ethical objections to the proposed research because (a) students', teachers', and administrators' time will be taken from potentially more beneficial educational experiences; (b) the study is likely to lead to unwarranted and inaccurate conclusions that may be damaging to the society that directly or indirectly pays for the research; and (c) the allocation of time and money to this poor-quality science will serve to keep those finite resources of time and money from better quality science. However, had the research question addressed been appropriate to the research design, these ethical issues would have been less acute. If the investigator had set out only to learn whether there were "performance differences between students in private versus public schools," the design would have avoided the causism problem and been appropriate to the question.

The analysis of research data is another area that raises ethical issues involving fidelity and scientific integrity. The most obvious and most serious transgression is the fabrication of data. Perhaps more frequent, however, is the omission of data contradicting the investigator's theory, prediction, or commitment. There is a venerable tradition in data analysis of dealing with outliers (extreme scores), a tradition going back over 200 years (Barnett & Lewis, 1978). Both technical and ethical issues are involved. The technical issues have to do with the best statistical ways of dealing with outliers without reference to the implications for the data analyst's theory (discussed in chapter 10). The ethical issues have to do with the relationship between the data analyst's theory and the choice of method for dealing with outliers. For example, there is some evidence that outliers are more likely to be rejected if they are bad for the data analyst's theory and are treated less harshly if they are good for the data analyst's theory (Rosenthal, 1978b; Rosenthal & Rubin, 1971). At the very least, when outliers are rejected, that fact should be reported. In addition, it would be useful to report in a footnote the results that would have been obtained had the outliers not been rejected.

Many researchers have been traditionally taught that it is technically improper (perhaps even immoral) to analyze and reanalyze their data in multiple ways (i.e., to "snoop around" in the data). We ourselves were taught to test the prediction with one particular preplanned analysis and take a result significant at the .05 level as our reward for a life well lived. Should the result not be significant at the .05 level, we were taught, we should bite our lips bravely, take our medicine, and definitely not look further at the data. Such a further look might turn up results significant at the .05 level, results

to which we were not entitled. All this makes for a lovely morality play, and it reminds us of Robert Frost's poem about losing forever the road not taken, but it makes for bad science and for bad ethics. It makes for bad science because, although exploratory data analysis does affect *p* values, it is likely to turn up something new, interesting, and important (Tukey, 1977). It makes for bad ethics because scientific data are expensive in terms of time, effort, money, and other resources, and because the antisnooping dogma is wasteful of time, effort, money, and other resources. If the research is worth doing, the data are worth a thorough analysis and being held up to the light in many different ways so that the research participants, funding agencies, science, and society will get their time and money's worth. We have more to say on this topic in chapter 10, but before leaving this issue, we should repeat that exploratory data analysis can indeed affect the *p* value obtained, depending on how the analysis was done. In chapter 14 we will show how statistical adjustments can be helpful here. Most important, replications will be needed no matter whether the data were snooped or not.

Although all misrepresentations of findings are damaging to the progress of science, some are more obviously unethical than others. The most blatant deliberate misrepresentation is the reporting of data that never were, which constitutes fraud (Broad & Wade, 1982). That behavior, if detected, ends (or ought to end) the scientific career of the perpetrator. **Plagiarism** (which comes from a Latin word meaning "kidnapper") is another breach of the fidelity principle; it refers to stealing another's ideas or work and passing it off as one's own—or as one author characterized it, "stealing into print" (LaFollette, 1992). A further distinction is sometimes made between "intentional" and "accidental" plagiarism (Rosnow & Rosnow, 2006). By *intentional plagiarism,* we mean the deliberate copying or taking of someone else's ideas or work and then knowingly failing to give credit or failing to place the quoted passage in quotation marks with a specific citation (Mallon, 1991). By *accidental plagiarism,* we mean the use of someone else's work but then innocently forgetting (not *neglecting*) to credit it (i.e., lazy writing). Intentional plagiarism is illegal, but this warning does not mean that researchers cannot use other people's ideas or work in their writing; it does mean that the writer must give the author of the material full credit for originality and not misrepresent (intentionally or accidentally) the material as one's own original work.

COSTS, UTILITIES, AND INSTITUTIONAL REVIEW BOARDS

We previously mentioned that, as required by federal law, institutions in which research with human subjects is conducted are required to maintain a review board (IRB) for the purpose of evaluating proposed investigations and monitoring ongoing research. The researcher provides the IRB with a detailed description (or "protocol") of the proposed investigation, and the IRB is then required (a) to evaluate whether the study complies with the standards for the ethical treatment of research participants and (b) to ensure that the potential benefits to individual participants (and society) will be greater than any risks the participants may encounter in the research (Stanley & Sieber, 1992). Some categories of studies may be exempt from IRB review, such as those in normal educational settings on normal educational processes; those involving educational tests, surveys, interviews, or observations of public behavior, as long

as the individual participants cannot be identified; and research involving existing public data (e.g., archival material) in which the individuals cannot be identified. In practice, however, universities often require that the protocol of any proposed study be submitted for review, so that the IRB can decide whether the study falls into a category that is exempt from review (Scott-Jones & Rosnow, 1998).

How can researchers forearm themselves against a capricious or overly zealous ethical review? There is no easy answer to this question, except to say that prudent researchers must sharpen their understanding of how risks and benefits are assessed in the review process (e.g., Brooks-Gunn & Rotheram-Borus, 1994; Ceci, Peters, & Plotkin, 1985; Diener & Crandall, 1978; Kimmel, 1996; Rosenthal, 1994c, 1995a; Rosnow, Rotheram-Borus, Ceci, Blanck, & Koocher, 1993; Wilcox & Gardner, 1993). When IRBs are confronted with a problematic or questionable protocol, they are expected to adopt a cost-utility approach in which the costs (or risks) of doing a study are evaluated simultaneously against such utilities (or benefits) as those accruing to the research participants, to society at large, and to science. Presumably, the potential benefits of higher quality studies and studies addressing more important topics are greater than the potential benefits of lower quality studies and studies addressing less important topics. Figure 3.1 shows a two-dimensional plane representing this type of analysis, in which the costs are one dimension and the utilities are the other (Rosenthal & Rosnow, 1984; Rosnow, 1990). In theory, any study with high utility and low cost should be carried out forthwith, and any study with low utility and high cost should not be carried out. Studies in which costs equal utilities are very difficult to decide on (B–C axis). In the case of low-cost, low-utility research, an IRB might be reluctant to approve a study that is harmless but is likely to yield little benefit.

As many researchers know from personal experience, however, the review process often ignores utilities and merely uses the A–C axis value for the criterion. Moreover, we have become convinced that, even when utilities are considered, this cost-utility model is insufficient because it ignores the costs of research not done. By concentrating only on the act of doing research and ignoring the act of not doing research, the review process uses a less rigorous standard of accountability than that aspired to by most researchers (Haywood, 1976; Rosenthal & Rosnow, 1984).

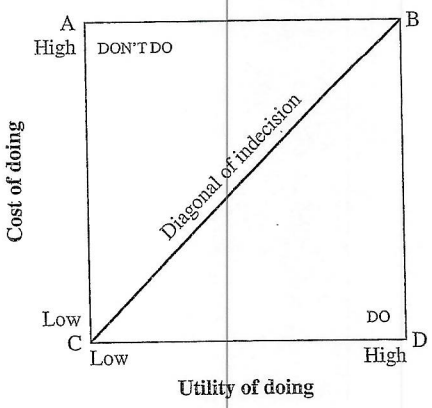

FIGURE 3.1
A decision-plane model of the costs and utilities of doing research. Studies falling at Point A are *not* carried out. Studies falling at Point D *are* carried out. Studies falling along the diagonal of indecision, B–C, are too hard to decide on.

Researchers often complain of their frustration at having research with possible societal benefits impeded by the review process or by political interference (Brooks-Gunn & Rotheram-Borus, 1994). In the 1990s, a prominent case involved a sexual survey of adolescents. The study was terminated prematurely on the grounds that it had violated community norms, but this decision simply deprived the community of data essential to addressing health problems of general concern (Wilcox & Gardner, 1993). If an IRB sent back a proposal for research that could conceivably find a way of preventing AIDS, on the grounds that its methodology did not ensure the privacy of the participants, the cost in human terms of the research not done could be high. Similarly, rejecting a sociopsychological investigation that might help to reduce violence or prejudice, but that involved a disguised experiment in a natural setting (i.e., a deception), would not solve the ethical problem, essentially trading one ethical issue for another.

It has been argued that it is incumbent upon researchers and their scientific organizations to educate IRBs about the costs of research not done, and about the costs to science and to society of not being able to replicate classic experiments that have generated important findings (Rosnow, Rotheram-Borus, Ceci, Blanck, & Koocher, 1993). A more complete analysis is represented by the two decision planes shown in Figure 3.2 or by the more complex model shown in Figure 3.3 (Rosenthal & Rosnow, 1984). Figure 3.2 is self-explanatory. Suppose, however, we added a new diagonal A–D (not shown) to these two planes and called it the "decision diagonal" (in contrast to B–C and B′–C′, the diagonals of indecision). For any point in the plane of *doing,* there would be a location on the cost axis and on the utility axis. Any such point could then be translated to an equivalent position on the decision diagonal. For example, if a point were twice as far from A as from D, we would see the translated point as located two thirds of the way on the decision diagonal A–D (i.e., closer to D than to A). The same reasoning would apply to *not doing,* except that closeness to A would mean "do" rather than "not do."

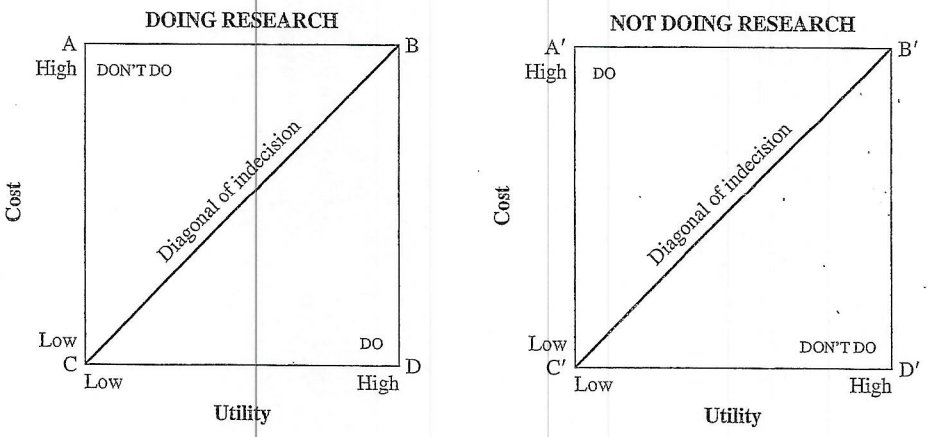

FIGURE 3.2
Decision planes representing the costs and utilities of doing (left plane) and not doing (right plane) research.

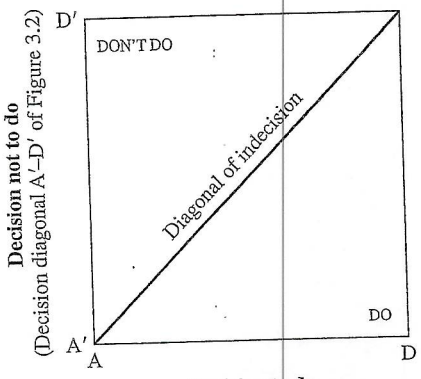

FIGURE 3.3
Composite plane representing both cases in Figure 3.2.

Putting these decision diagonals together gives Figure 3.3, and we are now back to two dimensions. In this composite plane, points near D tell us to do the study. Points near D′ tell us not to do the study. Points on the indecision diagonal leave us unsure. Points on the D′–D decision diagonal (not shown) tell us whether we are closer to "don't do" or "do." The purpose of this figure is to get us all to think about issues of cost and utility in terms of a more complete analysis. For example, the Tuskegee study reminds us that there have been shocking instances in which the safety of human subjects has been ignored or endangered (see also Beecher, 1970; Bok, 1978; Katz, 1972), but bureaucratic imperialism can also have serious consequences. As West and Gunn (1978) pointed out, if ethical guidelines are imposed absolutely, then "researchers may simply turn their attention to other topic areas that ethics committees and review boards find less objectionable" (p. 36). The result could be that research that needs to be done, to address vital scientific and societal questions, would cease. Considerations such as those indicated by Figures 3.2 and 3.3, if adopted by an IRB, would make it harder to give absolute answers to questions of whether or not particular studies should be carried out. Those who argue that a given study is unethical and should be prohibited would have to answer in ethical and moral terms for the consequences of *their* decision no less than the researchers proposing the study.

SCIENTIFIC AND SOCIETAL RESPONSIBILITIES

We have examined the kinds of questions concerning ethics and values that constantly confront behavioral researchers in studies with human participants. The ethical propriety of using animals in behavioral research has also attracted considerable attention, as research on animals has played a central role in our science since its beginning (e.g., studies by Ivan Pavlov and E. L. Thorndike). Attitudes toward the use of animals in behavioral research vary greatly among psychologists, most of whom nevertheless seem to approve of standard observational studies in which animals were confined in some way but apparently disapprove of studies involving pain or death (Plous, 1996).

Although the usual justification of animal research in psychology has been that it has clinical applications in humans and animals, questions have been raised about how often clinicians or clinical investigators actually draw on the results of animal research (Giannelli,1986; Kelly, 1986). On the other hand, a strong case has been made by many scientists that research on animals has been the foundation for numerous significant advances, including the rehabilitation of persons suffering from spinal cord injuries, the treatment of diseases and eating disorders, improvement in communication with the severely retarded, and a better understanding of alcoholism (Domjan & Purdy, 1995; N. E. Miller, 1985). For example, animal experiments by Roger Sperry, who won a Nobel Prize for his work, revealed that severing the fibers connecting the right and left hemispheres of the brain (resulting in a so-called split brain) did not impair a variety of functions, including learning and memory. That important discovery led to a treatment for severe epilepsy and made it possible for people who would have been confined to hospitals to lead normal lives (Gazzaniga & LeDoux, 1978; Sperry, 1968).

Just as the scientific community recognizes both an ethical and a scientific responsibility for the general welfare of human subjects, it also assumes responsibility for the humane care and treatment of animals used in research. There is an elaborate regulatory system to protect animals in scientific research, including professional and scientific guidelines (American Psychological Association, 1996), federal regulations, Public Health Service specifications, and state and local laws (see, e.g., Plous, 1996). Some years ago, the British zoologist William M. S. Russell and microbiologist Rex L. Burch made the argument that, given scientists' own interest in the humane treatment of the animals used in research, it would be prudent to search for ways to (a) *reduce* the number of animals used in research, (b) *refine* the experiments so that there is less suffering, and (c) *replace* animals with other procedures whenever possible. Called the "three Rs principle" by Russell and Burch (1959), this principle defines part of the moral contract to which researchers who use animal subjects subscribe.

Each researcher must weigh his or her responsibilities to science and to society very carefully. Even when the research is not directly funded by some agency of society, it is at least countenanced and indirectly supported, because our society places a high value on science and gives scientists a relatively free hand to study whatever they want to study. There are, to be sure, limits on how far scientists can go in the quest for knowledge, and we discussed some of those limits earlier. Society provides the circumstances and a psychological environment that are conducive to the practice of science. Because no scientist can guarantee that the outcome of his or her work will actually benefit society, what then does the scientist owe society in return for that privilege? As we have tried to show in this chapter, the researchers' ethical responsibilities are twofold: On the one hand, researchers must protect the integrity of their work in order to ensure that it measures up to the technical standards of sound scientific practice. On the other hand, researchers must also respect the dignity of those they study and the values that allow the pursuit of scientific knowledge in a free society.