Chapter 2 Reseach Methadology

ETHICS IN PSYCHOLOGICAL RESEARCH

PREVIEW & CHAPTER OBJECTIVES

This second chapter will introduce you to the most recent version (2002) of the ethics code formulated by the American Psychological Association (APA). The code directs psychological scientists in the planning, execution, and reporting of their research, and it includes guidelines for psychological research that tests both human participants and animals.1 Ethical issues are important to review early in this textbook because such issues must be addressed at all stages of the research process. When you finish this chapter, you should be able to:

Describe the origins and evolution of the APA ethics code.

Articulate the code's five general principles, especially as they apply to research in psychology.

Describe the role of the Institutional Review Board (IRB) in the research process and what needs to be done by the researcher to achieve IRB approval of research.

Explain when research proposals are exempt from IRB review, eligible for expedited review, or in need of a full formal review.

Explain why the decision-making processes of IRBs have occasionally been controversial.

Identify the essential features of a researcher's ethical responsibility when completing psychological research using adult human participants.

Describe historical examples of research that raised serious ethical questions.

Identify the ethical principles involved when completing research with children and those from special populations (e.g., prisoners and nursing home residents).

Describe how the ethics code applies to research that involves using the Internet.

Describe the arguments for and against the use of animals in psychological research.

Identify the essential features of a researcher's ethical responsibility when completing psychological research using animal subjects.

Identify the varieties of scientific dishonesty, how it can be detected, and understand some of the reasons why misconduct sometimes occurs in science.

A system of ethics is a set of "standards governing the conduct of a person or the members of a profession" (American Heritage Dictionary, 1992, p. 630). As members of the profession of psychology, researchers are obligated to follow the code of ethics established by the APA. When conducting research in psychology, our ethical obligations encompass several areas. Research psychologists must (a) treat human research participants with respect and in a way that maintains their rights and dignity, (b) care for the welfare of animals when they are the subjects of research, and (c) be scrupulously honest in the treatment of data. This chapter will examine each of these broad topics.

Before we describe the APA code of ethics, you should read Box 2.1, which describes one of psychology's best-known studies and two lesser-known experiments. The Little Albert experiment is often depicted as a pioneering investigation of how children develop fears, but it also serves well as a lesson in dubious ethical practice. Also, in the name of

psychological science, other infants have been subjected to repeated pinpricks in a study on adaptation to pain and have spent up to 14 months in relative isolation.

BOX 2.1 CLASSIC STUDIES—Infants at Risk

In this chapter, you will learn about an ethics code that is elaborate and finely tuned. In fact, you might think the code is unnecessarily complex and the good judgment of psychological researchers would surely prevent research participants from coming to serious harm. After you read about the following three studies, which occurred before the code existed, it should be clear why one was needed.

One of psychology's most frequently cited studies (Watson & Rayner, 1920) has come to be known as the Little Albert study. The authors were the famous behaviorist John B. Watson and Rosalie Rayner, a graduate student who eventually became Watson's second wife. The study tested just one child, an 11-month-old boy referred to as Albert B. The purpose of the study was to see if Albert could be conditioned to be afraid. Despite serious methodological weaknesses and failed replication attempts (Harris, 1979), the study has become a "classic" in psychology, routinely appearing in general psychology textbooks in the chapter on conditioning.

In prior research, Watson had determined that most infants were naturally afraid of loud noises and loss of support (e.g., falling). Watson and Rayner (1920) decided to use loud noise, produced when Watson struck a steel bar with a hammer just behind the infant's head. To see if the fear could be attached to a neutral stimulus, a white rat, the conditioning procedure was to pair the loud noise with the rat. When Albert reached out to touch the rat, "the bar was struck immediately behind his head" (p. 4). His response? "The infant jumped violently and fell forward, burying his face in the mattress" (p. 4). After several trials, the loud noise was no longer needed; Albert was now afraid of the rat. Because of generalization to similar stimuli, he was also fearful when shown a rabbit. Watson and Rayner made no attempt to remove the fear, although they made several suggestions for doing so.

It is difficult to hold Watson and Rayner responsible for ethical guidelines that were published several decades after they completed their study. Historical events must be evaluated in the context of their own times. They acknowledged, however, that "a certain responsibility attaches to such a procedure" (Watson & Rayner, 1920, p. 3). They decided to proceed because Albert seemed to be a strong, healthy child. Watson also justified the study by arguing that because Albert would learn such fears in real life anyway, he might as well learn them in a way that would advance behavioral science.

Watson and Rayner haven't been the only psychologists who used questionable judgment while studying infants. Two other examples are studies by Myrtle McGraw and by Wayne Dennis, both published in 1941. McGraw (1941) was interested in nervous system maturation, a legitimate topic of study. Her method was to apply repeated "pin pricks" to the cheeks, abdomens, arms, and legs of 75 children "at repeated intervals from birth to

four years" (p. 31). The pin pricks did not penetrate the skin, but they certainly caused distress, as is clear from McGraw's descriptions of the reactions to the stimulus. For example, she wrote that the "most characteristic response consists of diffuse bodily movements accompanied by crying, and possibly a local reflex withdrawal of the stimulated member" (p. 32). Eventually, just the mere sight of McGraw heading their way with a pin was enough to stress the children: "With advancing development, it will be observed that perception of the pin or of the approaching arm of the adult provokes fussing, crying, or withdrawal reactions on the part of this child" (p.33).

Dennis (1941) was interested in studying how early development would be affected by reducing environmental and social stimulation. From a local hospital, Dennis and his wife were able to "obtain" a pair of newborn female twins "because the mother was unable to provide for them" (p. 149). The Dennises offered the impoverished mother "temporary care of the twins in return for the privilege of studying them" (p. 149). The twins spent 14 months in the Dennis household, kept most of the time in a nursery room that afforded minimal views of the outside (sky and the top of a tree); the room contained little furniture and no toys. Dennis and his wife interacted with them only during feeding, bathing, and diaper changing, and "carefully refrained from rewarding or punishing the subjects for any action" (p. 150). Dennis reported delays in motor development for the girls but claimed no serious adverse effects resulted from the environmental deprivation. He concluded that during the first year, social interactions and environmental stimulation had minimal effect on children. He made little of the fact that the twins were slow in language development, an outcome that wouldn't surprise modern developmental psychologists. Today, research psychologists sometimes use animals in procedures that would not be considered appropriate for humans, and raising them briefly in isolation is an example. In 1941, however, Dennis had no misgivings about subjecting infants to an impoverished environment.

The Watson, McGraw, and Dennis studies were not completed by callous and unconcerned researchers but rather by people who believed they were advancing their science. But they were operating in the absence of a code of ethical conduct that might have given them pause. These studies make the need for an ethics code clear.

Developing a Code of Ethics for Psychological Science

After World War II, the United States and its allies conducted the Nuremberg trials to hold Nazi officers, doctors, and others associated with them accountable for various war atrocities. Included in those trials were the Doctors Trials which involved Nazi doctors' experimentation on human beings, including being injected with gasoline, deliberately infected with deadly diseases, and exposed to high levels of radiation for the purposes of sterilization and abortion. In addition, the implementation of the T-4 "Euthanasia" Program resulted in the systematic killing of individuals "unworthy of life" including children with physical or psychological disabilities between 1939 and 1945 (Weindling, 2004). One lasting legacy that emerged from these trials was the Nuremberg Code of ethics (1949), which emphasized the importance of voluntary consent from individuals

involved in medical research. You can review the complete Nuremberg code online at the Student Companion Site or by doing a simple Google search.

Psychologists in the United States published their first formal code of ethics in 1953 (APA, 1953), and it was influenced by the Nuremberg code. The document was the outcome of about 15 years of discussion within the APA, which had created a temporary committee on scientific and professional ethics in the late 1930s. This soon became a standing committee to investigate complaints of unethical behavior (usually concerned with the professional practice of psychology) that occasionally were brought to its attention. In 1948, this group recommended the creation of a formal code of ethics. As a result, the APA formed a Committee on Ethical Standards for Psychologists, chaired by Edward Tolman (Hobbs, 1948).

In keeping with psychology's penchant for drawing data-based conclusions, the APA committee took an empirical approach when developing the code. Using a procedure called the critical incidents technique, the committee surveyed the entire membership of the APA (about 7,500 members at the time), asking them to provide examples of "incidents" of unethical conduct they knew about firsthand and "to indicate what [they] perceived as being the ethical issue involved" (APA, 1953, p. 4). The request yielded over 1,000 replies. Although most concerned the practice of psychology (e.g., psychotherapy), some of the reported incidents involved the conduct of research (e.g., research participants not being treated well). A second committee, chaired by Nicholas Hobbs, then organized the replies into several drafts that were published in American Psychologist, APA's primary journal; readers were encouraged to comment on the drafts. The APA's council of directors accepted a final version of the code in 1952 and it was published the next year. Although it was concerned mainly with professional practice, one of its sections in this first ethics code was called "Ethical Standards in Research."

In the early 1960s, not long after APA created its first ethics code, a young, Yale psychologist named Stanley Milgram began a series of studies that became as well known for the questions they raised about research ethics as it did for their conclusions about human behavior. Milgram, who was Jewish, was motivated by questions about the Nazi Holocaust and deeply concerned about the problem of obedience to authority. (During the Nuremberg trials, a common defense used by Nazi war criminals was that they were just following orders.) Did the Holocaust reflect some basic flaw in the German psyche? Or is the tendency to obey authority found in all of us, produced when the circumstances are right? To answer his own questions, he developed his now famous research on obedience to authority. In the guise of a study on the effects of physical punishment on learning, Milgram induced volunteers to obey commands from an authority figure, the experimenter (who was actually a member of the research team and a high school biology teacher in real life). Playing the role of teachers, participants were told to deliver what they thought were high-voltage shocks (no shocks were actually given) to another apparent volunteer (also a member of the research team and a railroad payroll auditor in real life) who was trying, without much success, to accomplish a memory task (see Figure

2.1). A surprisingly high percentage of subjects complied with the “orders” from an experimenter to deliver shock and, in doing so, most subjects became quite distressed. In his original study, Milgram (1963) reported he had

observed a mature and initially poised businessman enter the laboratory smiling and confident. Within 20 minutes he was reduced to a twitching, stuttering wreck, and was rapidly approaching a point of nervous collapse. (p. 377)

FIGURE 2.1 The shock apparatus used by Milgram in his obedience studies.

As you might guess, Milgram’s research has been controversial. He was sharply criticized for exposing his volunteers to extreme levels of stress, for producing what could be long-term adverse effects on their self-esteem and dignity, and, because of the degree of deception involved, for destroying their trust in psychologists (Baumrind, 1964).

Milgram completed his obedience studies in the 1960s and early 1970s, which was a time in the United States when significant societal changes were taking place. The Civil Rights Act (1964) and the Voting Rights Act (1965) were passed by Congress and contributed to growing movement toward civil rights for all Americans. In part, the civil rights movement in American culture emphasized basic equality and dignity among individuals and created a heightened awareness of instances of inequality and mistreatment of vulnerable groups. One such group whose plight had finally come to light in the 1970s included poor Black men from the area around Tuskegee, Alabama, who were diagnosed with syphilis, but deliberately left untreated so that researchers could study the development of the disease over time (see Box 2.2). The revelation of the Tuskegee study in the early 1970s is one factor that led to the United States Congress to enact the National Research Act in 1974, which created the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. In 1979, the commission published what came to be called the Belmont Report, which includes three basic principles for research with human subjects: Respect for persons, Beneficence, and Justice. You can review the entirety of the Belmont Report online at the Student Companion Site or by doing a simple Google search.

Over the years, the APA ethics code has been revised several times, most recently in 2002. It currently includes a set of 5 general principles and 89 standards, the latter clustered into the 10 general categories. The general principles are “aspirational” in their intent, designed to “guide and inspire psychologists toward the very highest ideals of the profession” (APA, 2002, p. 1062), while the standards establish specific rules of conduct and provide the basis for any charges of unethical conduct.2 The entire APA code can be found online at www.apa.org/ethics/code/.

The five general principles reflect the philosophical basis for the code as a whole. These principles apply broadly to the science and practice of psychology. As they apply to research, they can be described as follows:

Beneficence and Nonmaleficence establishes the principle that psychologists must constantly weigh the benefits and the costs of the research they conduct and seek to achieve the greatest good in their research with little harm done to others.

Fidelity and Responsibility obligates researchers to be constantly aware of their responsibility to society and reminds them always to exemplify the highest standards of professional behavior in their role as researchers.

Integrity compels researchers to be scrupulously honest in all aspects of the research enterprise.

Justice obligates researchers to treat everyone involved in the research enterprise with fairness and to maintain a level of expertise that reduces the chances of their work showing any form of bias.

Respect for People's Rights and Dignity translates into a special need for research psychologists to be vigorous in their efforts to safeguard confidentiality and protect the rights of those volunteering as research participants.

There are many similarities between the principles of the Belmont Report and those of the APA Code, including some very similar language. Now, we turn to how psychological scientists apply these ethical principles in their research.

SELF TEST 2.1

How was the critical incidents technique used when the first APA ethics code was being developed?

What was the ethical justification used by Watson and Rayner in the "Little Albert" study?

What are the three basic principles of the Belmont Report?

The first general principle of the APA ethics code is "beneficence and nonmaleficence." What does this mean for the researcher?

Ethical Guidelines for Research with Humans

In the 1960s, a portion of the original ethics code was elaborated into a separate code of ethics designed for research with human participants. Another APA committee, led by Stuart Cook, used the same critical incidents procedure and published an ethics code specifically for researchers in 1973 (APA, 1973); it was revised in 1982 (APA, 1982) and again as part of the general revisions of 1992 and 2002. The specific APA Standards regarding research are found in Category 8 of the code ("Research and Publications"); you can find the full text of Category 8 at the Student Companion Site. In general, the standards for research with human participants include making a judgment that the benefits of the research outweigh the costs, gaining the informed consent of those participating in the study, and treating the research volunteers well during the course of the study and after it has been completed.3

Weighing Benefits and Costs: The Role of the IRB

All research on human behavior imposes some burden on those participating in the study. At a minimum, people are asked to spend time in an experiment when they could be doing something else. At the other extreme, they are sometimes placed in potentially harmful situations. In the name of psychological science, human research participants (or

subjects4) have received electrical shocks, been told they failed some apparently easy test, and been embarrassed in any number of ways. That such experiences can be distressing is illustrated in Milgram's (1963, 1974) obedience studies described earlier.

The basic dilemma faced by Milgram and every other researcher is to weigh the scientific value of the research being planned (a benefit) against the degree of intrusion on those contributing data to the study (a cost). On one hand, psychological scientists believe strongly in the need to conduct psychological research on a wide range of topics. Indeed, they believe that failing to investigate abdicates their responsibility as scientists. If the ultimate goal is to improve the human condition (the "Beneficence and Nonmaleficence" general principle), and if knowledge about behavior is essential for this to occur, then clearly it is essential to learn as much as possible. On the other hand, research can create discomfort for those participating in it, although few studies come anywhere near Milgram's experiments in terms of the level of stress experienced by subjects.

When planning a research study, the experimenter always faces the conflicting requirements of (a) producing meaningful research results that could ultimately increase our knowledge of behavior and add to the general good and (b) respecting the rights and welfare of the study's participants and causing them no harm. An integral part of the process of planning a study involves consulting with others. A good first step is to ask a researcher colleague whether your study has any ethical pitfalls. A formal process also exists, however, and it concerns a group called the Institutional Review Board or IRB. In a university or college setting, this group consists of at least five people, usually faculty members from several departments and including at least one member of the outside community and a minimum of one nonscientist (Department of Health and Human Services, 1983).5 In 1974, as part of the National Research Act, the federal government mandated that IRBs be in place for any college or university receiving federal funds for research. Today, IRBs are found in virtually all colleges and universities, whether or not federal funding is involved. Because of the complexity of the regulations involving research with human subjects, IRB members often go through a training program—a number of web-based programs exist (e.g., the Collaborative Institutional Training Initiative or CITI, found at www.citiprogram.org/). One survey found, however, that only 22% of IRB members reported having a formal training program at their institution (Cook & Hoas, 2011).

Researchers seeking IRB approval typically submit a rationale for the study and a description of research procedures, a statement about potential risks to participants, how these risks will be alleviated and why they can be justified, a copy of the study's informed consent form, and copies of materials to be used in the experiment. IRBs distinguish between proposals that are exempt from full review, those eligible for expedited review, and those requiring a full review. For research in psychology, proposals that are exempt from full review include studies conducted in an educational setting for training purposes (e.g., asking students like you to test each other on reaction time in the lab as part of a course requirement), purely naturalistic observation studies of public behavior, survey research that does not assess sensitive topics, and archival research. Proposals receiving expedited

review include many of the typical psychology laboratory experiments in basic processes such as memory, attention, or perception, in which participants will not experience uncomfortable levels of stress or have their behavior manipulated in any significant fashion. All other research usually requires a full review by the entire IRB committee.

As you might guess, there are gray areas concerning decisions about exempt, expedited, and full review. Hence, it is common practice for universities to ask that all research be given some degree of examination by the IRB. Sometimes, different members of an IRB are designated as "first step" decision makers; they identify those proposals that are exempt, grant approval (on behalf of the full board) for expedited proposals, and send on to the full board only those proposals in need of consideration by the entire group. At medium and large universities, where the number of proposals might overwhelm a single committee, departmental IRBs are sometimes created to handle the expedited reviews (Murphy, 1999).

An important component of an IRB's decision about a proposal involves determining the degree of risk to be encountered by participants. Sometimes, there is no risk at all, as when experimenters observe public behavior and do not intervene in any way. At other times, subjects in a study may be "at risk" or "at minimal risk." The distinction is not razor sharp but is based on the degree to which the people being studied find themselves in situations similar to "those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests" (Department of Health and Human Services, 1983, p. 297). Hence, subjects facing situations like those encountered in daily living that might involve some stress, but not a substantial amount, are considered to be "at minimal risk." If the risks, physical or mental, are greater than that, participants are said to be "at risk." For instance, people would be at minimal risk in a sports psychology study investigating whether training in visual imagery techniques leads to better athletic performance than the absence of such training. However, if that same study investigated whether the improvement due to training in imagery could be reduced by having participants ingest some drug, the degree of risk to participants would obviously be higher and require more careful scrutiny by an IRB.

When there is minimal or no risk, IRB approval is usually routinely granted through an expedited review, or the proposal will be judged exempt from review. However, when participants are "at risk," a full IRB review will occur and experimenters must convince the committee that (a) the value of the study outweighs the risk, (b) the study could not be completed in any other fashion, and (c) they will scrupulously follow the remaining ethical guidelines to ensure those contributing data are informed and well treated.

One final point about IRB approval is that when conducting research outside of the university environment, a researcher might have to satisfy more than a single review board. A health psychologist, for instance, might be using a local wellness center as a location for studying adherence to an exercise program. In addition to gaining university IRB approval, the

researcher will usually need an OK from the center's research committee before proceeding with the study.

IRBs provide an effective safeguard for participants, researchers, and universities, but they are controversial for three reasons. One issue is the extent to which IRBs should be judging the details of research procedures and designs (Kimmel, 2007). Researchers legitimately object to non-specialists (e.g., philosophy professors) passing judgment on methodologies they may not understand or research traditions they fail to appreciate. On the other hand, a poorly designed study has ethical implications. If it is seriously flawed methodologically, its results will be worthless, its participants could be harmed needlessly, and, at a minimum, their time will be wasted.

A second problem is that some researchers complain about IRBs being overzealous in their concern about risk, weighing it more heavily than warranted, relative to the scientific value of a study. For instance, a researcher described by Kimmel (2007) was unable to obtain IRB approval for a study in which people were asked to detect tones of varying loudness. Despite the fact that no tone was louder than conversational speech, the IRB insisted that listening to the tones "entailed a slight risk to [subjects'] welfare" (p. 283). The researcher refused to concede the point, argued with the IRB for 3 years, had no recourse for appeal, and eventually switched to animal research, stating that "the composition of animal welfare committees [was] a bit more reasonable" (p. 283). Obviously, not all IRBs are this arbitrary and inflexible, but the lack of an appeal process is a problem. Some studies have suggested that researchers, if they believe they have been treated unfairly by an IRB, might go as far as to omit from their IRB proposals some aspects of their procedure the IRB could find objectionable.

One unsettling consequence of IRBs being overly conservative, according to prominent social psychologist Roy Baumeister, is that psychology is rapidly becoming the science of self-reports and finger movements (keystrokes on a computer) instead of the science of overt behavior. After examining recent issues of the Journal of Personality and Social Psychology, Baumeister and his colleagues discovered that the "[d]irect observation of meaningful behavior is apparently passé" (Baumeister, Vohs, & Funder, 2007, p. 397). Instead, it seemed that in the articles they read, subjects spent most of their time filling out surveys or describing how they or others might behave in some hypothetical situation. One explanation for the shift from overt behavior to self-report studies is efficiency; studies that ask subjects to read a scenario (e.g., encountering someone in need of help) and predict how they or others would react can be completed much more quickly than studies that actually place subjects in that scenario and record how they actually react. But another reason, Baumeister et al. argued, has to do with getting IRB approval. Measuring meaningful social behavior (as in the helping behavior example) usually means using deception, and it therefore places more of a burden on researchers to show their participants will be protected. Self-report studies are safer. Although Baumeister et al. acknowledged that important things about people can be learned from self-reports, he

worried that psychology was showing signs of "abandoning its original goal of being the science of behavior" (p. 400).

A third issue that concerns psychologists is that IRBs sometimes overemphasize a biomedical research model to evaluate proposals. As a result, they might ask researchers to respond to requests that are not relevant for most psychological research. For example, they might ask that the consent form include information about procedures or alternative courses of treatment available to those who choose not to participate in the study (Azar, 2002). This makes sense for research evaluating the effectiveness of some medical treatment but makes no sense in most psychological research, where the alternative to participating is simply not to participate. Susan Fiske (2009), a prominent social psychologist and former chair of the IRB at Princeton University, recommended that universities sponsoring medical research should create separate IRBs for medical and behavioral research.

One unfortunate consequence of these three issues is a lack of consistency among IRBs. Several studies have shown that identical IRB proposals have fared differently with different IRB committee members. In one example, researchers proposed a study in which 6- to 10-year old children would view a 4-minute video in which a child actor mistakenly claimed to have been hit by a firefighter (who had in fact just told the child actor to leave his firefighter's hat alone). Children viewing the video would then be interviewed to determine their ideas about why the child actor might have lied. An IRB rejected the proposal on the grounds that "it was deemed unethical to show children public servants in a negative light" (Ceci & Bruck, 2009, p. 28). The identical IRB proposal, however, had been approved by IRBs at two other universities, had been found ethically acceptable by the National Science Foundation (which was funding the research), and was judged harmless by a panel of pediatricians and child development specialists.

Despite these issues, the primary goal of IRBs is to evaluate any ethical concerns that may arise during the course of the proposed research study. Because IRBs are comprised diverse members and may not include psychologists, they are often guided by the Nuremberg code or Belmont Report, which are broader in scope than the APA code. However, it is the responsibility of the psychology researcher to adhere to the APA code when proposing to the IRB psychological research involving humans because the APA code both encapsulates the Nuremberg code and Belmont Report as well as expands upon them.

Informed Consent and Deception in Research

A central feature of the APA code is the concept of informed consent (Standard 8.02), the notion that in deciding whether to participate in psychological research, human participants should be given enough information about the study's purpose and procedures to decide if they wish to volunteer. For example, the use of painful procedures in a study (e.g., electric shock and regardless of how mild it is) must be disclosed. Consent procedures evolved from the aftermath of historical abuses, most notably the medical research conducted in Germany during World War II that used concentration camp

inmates as human guinea pigs. At the Nuremberg trials, the Nazi doctors defended their actions by arguing that voluntary consent didn't really exist in any medical research of the time and that the long-term importance of their research outweighed any adverse consequences to the participants. Their argument failed, they were convicted, and the presiding tribunal wrote what was called the Nuremberg Code mentioned earlier. It established the principle that consent must be informed, competent, and voluntary and that the person giving it must be able to comprehend the situation involved (Faden & Beauchamp, 1986).

Although the experiments performed on concentration camp victims are the most dramatic and appalling examples of consent violations, problems have occurred in the United States as well. See Box 2.2 for brief descriptions of cases in which (a) children with severe intellectual disabilities were infected with hepatitis in order to study the development of the illness; (b) southern Black men with syphilis were left untreated for years and misinformed about their health, also for the purpose of learning more about the time course of the disease; and (c) Americans, usually soldiers, were given LSD without their knowledge.

BOX 2.2 ETHICS—Historical Problems with Informed Consent

The research activities of doctors in the Third Reich are unprecedented in their callousness and cruelty. Nonetheless, there are cases in the United States of research projects that have provoked intensely critical reactions and have invited comparisons, albeit remote, to the Nazi doctors. Three famous examples are the Willowbrook hepatitis study, the Tuskegee syphilis study, and project MK-ULTRA.

At Willowbrook, an institution housing children with varying degrees of mental disability, an experiment began in 1956 and continued into the 1970s in which approximately 1 in 10 new admissions was purposely infected with hepatitis. The parents were told of the procedure and agreed to it, but it was later shown that they might have felt pressured into giving consent. The Willowbrook study was investigating hepatitis, not mental disability.

The study was initiated because hepatitis was rampant at Willowbrook institution, partly due to a high proportion of severely disabled children who could not be toilet-trained. At one point in the 1950s, there were 5,200 residents; of those, 3,800 had IQs lower than 20 and more than 3,000 were not toilet trained (Beauchamp & Childress, 1979). Even with the staff's best efforts, conditions were generally unsanitary and led to the spread of the disease. By deliberately infecting new admissions and placing them in a separate ward but not treating them, the researchers hoped to study the development of the disease under controlled conditions. Those in charge of the project defended it on the grounds that the children would almost certainly contract the disease anyway, so why not have them contract it in such a way that more could be learned about it? Indeed, although the study has been legitimately criticized on ethical grounds, it did contribute to the understanding of hepatitis and improved treatment of the disease.

The Tuskegee study was designed to examine the physical deterioration of persons suffering from advanced syphilis (Jones, 1981). Beginning in the early 1930s, about 400 poor Black men from the rural South were diagnosed with the disease and deliberately left untreated. They were never informed about the nature of the disease, nor were they told its name; doctors simply informed them they had "bad blood." Also, local physicians agreed not to treat the men. Given the poverty of the participants, it was not difficult to induce (coerce?) them to visit the clinic periodically (free rides and a hot meal), where blood tests and other examinations were done. The project continued into the early 1970s, even though it was clear by the late 1940s that the subjects were dying at twice the rate of a control group and were developing significantly more medical complications (Faden & Beauchamp, 1986). Defenders of the study argued that, when it began in the 1930s, there was no effective treatment for the disease and little knowledge of it. Like Willowbrook, the Tuskegee study contributed to our understanding of a serious disease, but its value was vastly overshadowed by the consent violations.

While the chief investigators in both the Willowbrook and Tuskegee studies were misguided in their abuse of the informed consent concept, they had a strong desire to learn as much as possible about two devastating diseases, hepatitis and syphilis. The third example of a consent violation, unfortunately, lacked even the justification of an eventual medical benefit. This was a project launched by the Central Intelligence Agency (CIA) to expose unknowing human participants to the drug LSD in order to gauge the drug's ability as a weapon. The project was created in the early 1950s, during the Cold War between the United States and the former Soviet Union. Prompted by an erroneous intelligence report that the Soviets were buying up the world's supply of LSD (Thomas, 1995), CIA leadership approved a program to determine if LSD could cause mental confusion or render captured spies defenseless. Over approximately 10 years, the CIA sponsored numerous studies on unwitting participants, often soldiers but sometimes members of the general public. Soldiers signed consent forms, but the forms said nothing about the potential effects of the drug and were designed mostly to ensure that soldiers would not reveal their participation. That secrecy was important is clear from an internal CIA memo that read, in part,

Precautions must be taken … to conceal these activities from the American public. … The knowledge that the Agency is engaging in unethical and illicit activities would have serious repercussions in political and diplomatic circles. (cited in Grose, 1994, p. 393)

What went on during MK-ULTRA? Projects included giving soldiers LSD and then putting them in isolation, giving them the drug and then performing a lie detection task, examining the effects of repeated doses (over 77 consecutive days in one case), and even surreptitiously giving the drug to men visiting prostitutes in a CIA-financed brothel, with agents observing behind two-way mirrors (Thomas, 1995). This latter study was code-named by a CIA humorist—it was called "Operation Midnight Climax."

At least two people died as part of MK-ULTRA and numerous others were adversely affected by it. Consider this case, as described in the 1994 Rockefeller Report, the results of a congressional investigation into 50 years of CIA-sponsored biological experimentation:

In 1957, _____ volunteered for a special program to test new military protective clothing. He was offered various incentives to participate in the program, including a liberal leave policy. … During the 3 weeks of testing new clothing, he was given two or three water-size glasses of a liquid containing LSD to drink. Thereafter, Mr. --- developed erratic behavior and even attempted suicide. He did not learn that he had received LSD … until 18 years later, as a result of congressional hearings in 1975. (Rockefeller Report, 1994)

The CIA did not bother to inform either Congress or the President about MK-ULTRA. The program ground to a halt in 1963, not because of any ethical misgivings on the part of the CIA, but primarily because the studies had not yielded any useful military information. Congressional investigators discovered it in the mid-1970s and eventually issued a full report (Grose, 1994).

Typical consent forms contain several features. First, potential volunteers agree to participate after learning the general purpose of the study (but not the specific hypotheses), the basic procedure, and the amount of time needed for the session. Second, participants understand they can leave the session at any time without penalty and with no pressure to continue. Milgram's subjects were encouraged to continue by the experimenter, as this was part of Milgram's procedure. In spite of subjects' lack of willingness to continue the experiment, Milgram's experimenters were trained to say thing like "The experiment requires that you continue" and "It is absolutely essential that you continue" (Milgram, 1963, p. 374). This clearly violates participants' freedom to discontinue a study at any time.6 A third major feature of consent forms is that participants are informed that strict confidentiality and anonymity will be upheld. This feature is closely related to APA General Principle E described earlier. Fourth, if questions linger about the study or if they wish to complain about their treatment as participants, there are specific people to contact, including someone from the IRB. Finally, participants are informed of any risk that might be encountered in the study, and they are given the opportunity to receive a summary of the results of the study, once it has been completed. When writing a consent form, researchers try to avoid jargon, with the aim of making the form as easy to understand as possible. Examples of a consent forms are often provided by the institution's IRB. We encourage you to explore the website for your institution's IRB to familiarize yourself with its code of ethics, committee membership, instructions for submission, and sample IRB forms. We have provided a sample of a consent form online at the Student Companion Site

A new feature of the 2002 revision of the ethics code is a more detailed set of provisions for research designed to test the effectiveness of a treatment program that might provide benefits but might also be ineffective and perhaps even harmful (Smith, 2003)—a program to treat post-traumatic stress disorder, for instance. This revision is found in

Standard 8.02b, which tells researchers to be sure to inform participants that the treatment is experimental (i.e., not shown to be effective yet), that some specific services will be available to the control group at the end of the study, and that services will be available to participants who exercise their right to withdraw from the study or who choose not to participate after reading the consent form. Participants must also be informed of the method by which people have been assigned to the treatment and control groups.

Although informed consent is essential in most research in psychology, it is important to note that consent is not required for research that is exempt from full review. As Standard 8.05 indicates, consent is not needed in studies using anonymous questionnaires, for data that have already been collected for another purpose (archival data), for classroom projects in which data collection is for demonstration purposes, and for certain employment-related data collection exercises. Also, consent is not needed for observational studies that occur in certain locations; the key is whether the setting is a public one—if the study occurs in a place where anyone could be observed by anyone else, consent is not needed (Koocher & Keith-Spiegel, 1998).

Informed Consent and Special Populations

Not all research participants are capable of giving consent, due to factors as age or disability, and some persons might experience undue coercion to volunteer for research (e.g., prisoners). In these circumstances, additional procedures apply. For example, the Society for Research in Child Development (SRCD) follows a set of guidelines that expand upon some of the provisions of the code for adults. Thus, because children (anyone under age 18) might not be able to fully understand consent forms, their parents or legal guardians are the ones who give consent. Nonetheless, unless the participant is an infant or is otherwise not capable of skilled language use, researchers are obligated to inform the child about the study and to gain what is referred to as assent. That is, researchers give the child as much information as possible to gauge whether the child is willing to participate. According to the SRCD code, assent occurs when "the child shows some form of agreement to participate without necessarily comprehending the full significance of the research necessary to give informed consent" (SRCD, 1996, p. 337). Assent also means the researcher has a responsibility to monitor experiments with children and to stop them if it appears that undue stress is being experienced. A parent may give informed consent for a study on the effects of TV violence on children's aggressive behavior, but the parent might not be in the room when the film is shown. It is up to the researcher to be sensitive enough to remove the child from the task at hand (and repair the damage) if the stress level is too high.

In addition to the assent provision, the SRCD code requires that additional consent be obtained from others who might be involved with the study in any way. For example, this would include teachers when a study includes their students. The code also cautions researchers about incentives that might be used, either to induce a willingness to participate or as rewards for tasks completed. The rewards "must not unduly exceed the

range of incentives that the child normally receives" (SRCD, 1996, p. 337). Also, researchers should not use the potential rewards as an inducement to gain the child's assent; indeed, rewards should not even be mentioned until after the parents have given full informed consent (Scott-Jones, 2000). Finally, the SRCD code mirrors the provisions of the code for adults, but warns researchers to be even more vigilant in certain areas. These include the decisions about balancing scientific gain against risk to participants, the level of deception that can be justified, and the reporting of the study's results.

Additional provisions for the protection of participants exist with other special populations. Thus, legal guardians must give truly informed consent for research with people who are confined to institutions (e.g., the Willowbrook case). Second, it is imperative to ensure that participants do not feel coerced into volunteering for a study. This problem is difficult to avoid in environments such as prisons because even with the best intentions of researchers, prisoners might believe that their failure to volunteer will cost them in the future and perhaps even affect their future parole status. In general, researchers tend to rely on simple material rewards (e.g., money) and make it clear to prisoners that their participation will not be noted in any way in their parole records (Diener & Crandall, 1978). As was the case for the SRCD code for research with children, the inducements to participate must be reasonable.

Another issue with confined populations is confidentiality (Kimmel, 2007). While normal guidelines for disguising the identity of participants apply, researchers are legally obligated to break confidentiality under circumstances that involve a clear danger (e.g., a prisoner participant reveals he is about to kill another prisoner). Finally, as illustrated in Box 2.2 in the Willowbrook case, research with confined populations should be designed for the expressed purpose of providing knowledge that will in some way benefit the members of that population.

Use of Deception

Consider the following scenario: You decide to sign up for an interesting-looking psychology experiment on problem solving. You show up at the appropriate time and place and, after being given initial instructions by an experimenter, you and another participant are left alone and given some anagrams to solve (anagrams are sets of letters that have to be unscrambled to make a word). After 5 minutes or so, the other person seems to get upset about the difficulty of the task and then storms out of the room. The experimenter returns, asks you a series of identifying questions about the person who just left (e.g., "Could you describe what she was wearing?"), and then asks you to identify this person from a set of photos. The experimenter then informs you that the real purpose of the study was eyewitness identification accuracy, not anagram problem solving. How would you react to this?

Standard 8.07 of the APA code indicates subjects might experience deception in a study if it is determined by the researcher, and agreed to by the IRB, that the study could not be done in any other fashion. That is, participants might not be told the complete details of a study

at its outset, or they might be misled about some of the procedures or about the study's purpose, as in the eyewitness example you just read. Researchers argue that in the absence of deception in certain studies, participants would not act naturally. If you knew you were in a study on eyewitness identification and that the anagrams didn't matter, you probably wouldn't bother much with the anagrams. Instead, you'd be trying to memorize the features of the other person in the room, a behavior that would not occur in a real-world eyewitness situation. How can these apparently contradictory concepts of consent and deception be reconciled?

One could argue that truly informed consent should never result in people being deceived about the purposes of the study. Some (e.g., Baumrind, 1985) have recommended eliminating deception in all psychology experiments on the grounds that people in positions of trust (i.e., experimenters) should not be lying to others (i.e., subjects). The outcome of deceptive research, she believes, is that participants could become mistrustful of experts and perhaps even cynical about the legitimacy of psychology as a science. Others (e.g., Geller, 1982) have argued that the need for "truth in advertising" could be met by forewarning those thinking about participating in a study that involves deception. They could be given a general rationale for deception during the consent procedure, told that some form of deception would probably occur in the study, and assured that all would be revealed at the end. Forewarning has been criticized, however, on the grounds that subjects would spend more time trying to figure out the true purpose of the study than they would behaving naturally and that many would refuse to participate, thereby reducing the accuracy of the study's results (Resnick & Schwartz, 1973).

Milgram's (1963, 1974) obedience studies provide a further illustration of why psychologists sometimes withhold information about the true purpose of the study at the beginning of the experiment. We've seen that Milgram told his subjects he was investigating the effects of punishment on learning. Teachers (the real subjects) tried to teach a list of word pairs to the learner, believing they were shocking him for errors. Milgram was not really interested in learning, of course. Rather, he wanted to know whether his volunteers would (a) continue to administer apparent shocks of increasing voltage to a learner who was in discomfort and not learning much, or (b) disobey the experimenter and stop the experiment. The outcome: Few people disobeyed. In the original study, 26 out of 40 continued shocking the learner even when the voltage level seemed to reach 450 and nobody disobeyed until the level reached 300 volts (Milgram, 1963)! If Milgram had informed his "teachers" he was interested in seeing whether they would obey unreasonable commands, would the same results have occurred? Certainly not. Blind obedience to authority is not something people value highly, so subjects told ahead of time they are in a study of obedience would surely be less compliant than they otherwise might be. The point is that researchers want their participants to take the task seriously, to be thoroughly involved in the study, and to behave as naturally as possible. For that to happen, deception is sometimes necessary. Please keep in mind, however, that the Milgram study is an extreme example of deception. Although deception studies with elaborate cover stories are more likely to be found in social psychology than in other

research areas (Korn, 1997), the level of deception is minor in most research. Typically, it involves the withholding of some information about the study rather than a cover story that creates the impression that the study concerns topic A when it really involves topic B. That is, most deception research involves omitting some information in the consent process rather than actively misleading participants about what they are to encounter (Fischman, 2000). For instance, participants in a memory study might be given a series of five word lists to study and recall, one at a time. At the end of the session, although not initially informed of it, they might be asked to recall as many words as they could from all five lists. Information about that final recall would be omitted from the original instructions to get a better measure of the memory for all of the lists, uncontaminated by extra rehearsal.

Treating Participants Well

Several portions of the ethics code are designed to ensure that volunteers are treated fairly and with respect during their participation, that they receive complete information about the study at its conclusion, that any stress they encounter is relieved, and that their participation is kept private. It is important to note this responsibility extends to everyone involved in the running of the study, from the primary researcher to the graduate students or undergraduates who might actually run the experimental sessions.

We have already seen the researcher must estimate the amount of risk to participants, with greater amounts of risk creating a greater burden to justify the study. This problem of risk and potential harm is addressed in the standards relating to informed consent and use of deception and once more in Standard 8.08, which makes it clear that responsibility does not end with the conclusion of the testing session. After the study is over, the researcher has an additional task, called debriefing, during which the researcher answers questions the participants might have and tells them about the purpose(s) of the study. It is not essential that participants be informed about all aspects of the study immediately after their participation. Standard 8.08(b) "[I]f scientific or humane values justify delaying or withholding this information, psychologists take reasonable steps to reduce the harm" makes it clear that, in some circumstances, the immediate debriefing can be incomplete. This situation occurs most frequently when some deception is involved, college students are the population under study, and the experimenter is concerned about participants talking to other potential participants (classmates). This latter problem, sometimes referred to as participant crosstalk, can ruin a study. Even in a study with relatively minor deception, subjects who go into the study knowing something unknown to naïve subjects will certainly be influenced by their knowledge of the study.

There is evidence that participant crosstalk occurs, especially in situations where participants (e.g., college students) can easily interact with each other (Diener, Matthews, & Smith, 1972). A recent study confirmed the problem still exists by cleverly determining the frequency of its occurrence. Edlund, Sagarin, Skowronski, Johnson, and Kutter (2009) had subjects estimate the number of beans in a jar. Those participating were then given the correct answer. The question was whether or not these subjects would pass the

information along to future participants. Some clearly did just that, although the percentage doing so was small (just under 5%). The percentage was reduced in a second study, when participants were specifically asked not to reveal the number of beans to others who might participate. Aside from urging subjects not to discuss the study after their participation, a common strategy for reducing crosstalk, consistent with Standard 8.08(b), is to provide information about the general nature of the research during debriefing but to provide full information about the study only after the experiment has been completed.

In general, debriefing serves two related purposes, referred to by Holmes (1976a, 1976b) as dehoaxing and desensitizing. Dehoaxing means revealing to participants the true purpose of the study and the hypotheses being tested (or some portion of them), and desensitizing refers to the process of reducing stress or other negative feelings that might have been experienced during participation in the study. Subjects are also informed that, if they wish, they may have their data removed from the data set.

The amount of time spent in debriefing depends on the complexity of the study, the presence and degree of deception, and the level of potential distress. In a study involving deception, the researcher often begins a debriefing session by asking participants if they thought the study had a purpose other than the one initially described. This enables the researcher to determine if the deception was effective; it also provides a lead-in to further explication of the study. At this time, the researcher tries to justify the deception (e.g., emphasizes the importance of getting one's true reactions) and begins to alleviate stress. Participants taken in by the experiment's cover story are told their behavior reflects the effectiveness of the cover story, not any personal weakness. That is, subjects in many types of studies can be assured that the situation they experienced had powerful effects on their behavior, that their reactions don't reflect any individual inadequacies, and that others reacted similarly (Holmes, 1976b). In most cases, dehoaxing amounts to explaining the importance of eliciting natural behaviors and discussing the nature of the research topic being studied.

Several studies have shown that participants who are thoroughly debriefed evaluate the research experience positively. Smith and Richardson (1983) showed that, compared to nondeceived subjects, those in deception studies actually rated their experiences higher in both enjoyment and educational value, apparently because the debriefing was more extensive. One result of an effective debriefing is that skilled researchers can better understand their current study and improve future ones. Participants can be asked for their ideas about revising the procedure in order to learn more about the problem being studied. In many cases, their descriptions of what they were thinking about during the experiment can be helpful in interpreting the data and planning the next study.

The importance of leaving people with a good feeling about their research participation cannot be overstated. Yet it can be a difficult business, especially when deception is involved. Consider the Milgram experiment again: What must that debriefing have been

like? In fairness to Milgram, he was apparently sensitive to the emotional health of his subjects. After the study was completed, he sent them a survey about their experience and a five-page report describing the results and their significance. The results of the survey indicated that 84% of participants stated they were glad to have participated (Milgram, 1964). He also completed a 1-year follow-up study in which a psychiatrist examined 40 former participants and found "no evidence … of any traumatic reactions" (Milgram, 1974, p. 197).

Studies surveying research volunteers have found that fears of excessive harm in psychological research might be exaggerated; participants seem to understand and accept the rationale for deception (Christensen, 1988). One survey even found that college students were considerably more lenient than professional psychologists in their judgments about the ethical appropriateness of four hypothetical studies involving such things as experimentally produced stress and alterations of self-esteem (Sullivan & Deiker, 1973). Other research shows objections by subjects to participating in psychological research seem to center more on their concern about being bored than being harmed (Coulter, 1986). On the other hand, it has been argued that post-experiment surveys of participants are biased, especially if deception has been involved. Having been misled and perhaps embarrassed in the study, deceived participants might respond positively to surveys as part of the process of convincing themselves the study was worth their time and effort (Baumrind, 1985)—another example of an effort justification (see Chapter 1). This phenomenon probably accounted for at least some of Milgram's survey results (the 84%). Fisher and Fyrberg (1994) avoided this post-deception survey problem by asking students who had not yet been participants in research to evaluate three published studies involving various forms of deception. They found students believed that participants would be embarrassed or made uncomfortable in the studies and that debriefing, while essential, would not completely alleviate the negative feelings. Yet, when asked to make an overall cost-benefit assessment of the three studies described, 90%, 73%, and 79% (depending on the described study) of the students judged the scientific merit of the research to be sufficient to justify the deceptions involved.

One last aspect of treating participants well concerns privacy and confidentiality, which is encapsulated by APA General Principle E. Research participants should be confident their identities will not be known by anyone other than the experimenter and that only group or disguised (coded) data will be reported. The only exceptions to this occur in cases when researchers might be compelled by law to report certain things disclosed by participants (e.g., child abuse, clear intent to harm oneself or another). In research that could involve such disclosure, researchers should word the consent form to make it clear that confidentiality could be limited (Folkman, 2000). The basic right to privacy also applies to research outside of the laboratory that might affect people in daily living situations. As we'll see in the next chapter, when laboratory and field research are compared, concerns over invading the privacy of people going about their daily business keeps many researchers within the protected confines of the laboratory.

In summary, in research using human participants, our ethical obligations under the APA code include the following:

Developing a study in which the overall benefits outweigh the overall costs
Avoiding doing anything that would harm participants
Gaining informed consent (under most circumstances)
Assuring volunteers they can quit the study at any time, without penalty
Providing some form of debriefing
Assuring participants about confidentiality and their anonymity
Research Ethics and the Internet

Because the Internet has altered life dramatically in the 21st century, you won't be surprised to learn that research in psychology has been affected by the electronic world. Online research methods of interest to psychologists falls into two broad categories (Anderson & Kanuka, 2003). First, some websites are designed to collect data from those logging into the sites. This happens most frequently in the form of online surveys and questionnaires but can involve other forms of data collection as well. For example, Amazon's Mechanical Turk (MTurk: www.mturk.com) allows researchers to generate surveys or program experiments, solicit participation, and pay subjects for their participation at very low cost. In other cases, subjects login to sites controlled by researchers on their own campus, and complete a study electronically (e.g., a survey created on software such as Survey Monkey or Qualtrics). Some research has been completed on the issue, and it appears that data collected electronically correspond reasonably well and yield similar results as data collected in a more traditional fashion (Casler, Bickel, & Hackett, 2013; McGraw, Tew, & Williams, 2000). However, it is important to also note that there are differences in the characteristics of MTurk users versus participants typically tested face-to-face. In addition to being more socially and economically diverse, MTurk users are also more likely to use the Internet to look up answers to factual questions, tend to have lower self-esteem, and be more introverted, posing challenges to research where such factors may be relevant (Goodman, Cryder, & Cheema, 2013).

The second form of online research involves a researcher studying the behavior of Internet users. This research ranges from examining the frequency of usage of selected websites to analyses of the content of web-based interactions (e.g., monitoring the activity of a Twitter feed). For both types of research, the basic principles of the ethics code apply, but research involving the Internet introduces unique ethical problems for the researcher. The problems have even resulted in the development of a code of ethics for Internet-based research, created by an organization called the Association of Internet Researchers. The American Association for the Advancement of Science has also prepared guidelines for IRBs that must decide whether to approve online research, and the APA's Board of Scientific Affairs established an advisory group on the Internet in 2001 and published its report 3 years later (Kraut, Olson, Banaji, Bruckman, Cohen, & Couper, 2004).

For online research in which computer users contribute data, problems relating to informed consent and debriefing exist. During a normal informed consent procedure, the

experimenter can quickly clear up any confusion or misunderstanding on the part of participants and can be reasonably sure participants read the consent form before signing. Consent forms can be used easily enough in online studies, but there is no opportunity for researchers to answer questions (although some consent forms are accompanied by a set of frequently asked questions and their answers) and no way to know if the consent form has been read. Another consent problem concerns age: Researchers can post warnings that participants need parental consent if they are under age 18, but it is impossible to monitor compliance. Debriefing may also be problematic. A good debriefing session is interactive, with questions asked and answered, but with online research, there is no guarantee participants will even be there to read the debriefing information. One click and the participant is gone without being debriefed. Furthermore, if deception is involved, while the dehoaxing part of debriefing can be managed by presenting clear information, the desensitizing part will be difficult if not impossible to accomplish.

Online research involving the collection of information from computer users involves an additional set of problems. A major issue concerns privacy and confidentiality (Kraut et al., 2004). As you recall from the discussion of informed consent earlier in this chapter, consent is not required for studies that are purely observational and individual behavior is observed in public places. With online research, the interesting and as yet unresolved question is whether activities such as Twitter feeds, Facebook posts, chat rooms, blogs, discussion boards, and listservs are public forums or private discussions. For the researcher, the best guideline is be faithful to the general principles of the code and to consult frequently with colleagues and the local IRB during the planning stages of online research. For users of social media, it is important to be aware that, in the absence of sophisticated encryption software, messages posted are "out there," available to anyone with an Internet connection. The best advice for users is to think of the messages they post as having about the same level of privacy as postcards.

Concerning confidentiality, researchers using Internet surveys must take steps to ensure the protection of the user's identity even if the participant used her or his own personal computer. This can mean ensuring that cookies (tools used to track information about Internet users) are not left on the participant's computer as a result of taking the survey. Also, users must be assured that if their computer's identity (e.g., an IP address) is returned with the survey, the researcher will discard the information (Pollick, 2007).

SELF TEST 2.2

You wish to do a study comparing two memory improvement techniques. Which category of the IRB approval process will apply in this case?

How does the APA define informed consent?

Milgram's procedure probably would not have gained IRB approval in terms of its consent procedures. What was the most obvious problem?

Ethical Guidelines for Research with Animals

As you recall from your course in introductory psychology, psychologists occasionally use animals as research subjects. Although some people have the impression that

psychologists study rats more than people, the truth is that animal research involves a relatively small proportion of the total research done in psychology, about 7–9% (Gallup & Suarez, 1985a). Also, the vast majority of studies use rats and mice as subjects; dogs, cats, and nonhuman primates are used in just a tiny proportion of animal research. Despite the small proportions, many of psychology's important contributions to human welfare are based on a foundation of research with animals (Domjan & Purdy, 1995).

Animals are used in psychological research for several reasons. Methodologically, their environmental, genetic, and developmental histories can be easily controlled. Genetic and life-span developmental studies can take place quickly—female mice, for instance, produce litters after just 3 weeks of pregnancy, and 1 mouse year is the equivalent of 30 human years (Herzog, 2010). Ethically, most experimental psychologists take the position that, with certain safeguards in place, animals can be subjected to procedures that could not be used with humans. Consider Eleanor Gibson's visual cliff research again (Gibson & Walk, 1960; see Chapter 1). Thirty-six 6- to 14-month-old infants were placed in the middle of the apparatus, and although they were quite willing to crawl around on the "shallow" side, they hesitated to crawl onto the glass surface over the "deep" side. This shows that they were able to perceive depth and apparently were aware of some of its consequences. Does this mean depth perception is innate? No, because these infants had 6 to 14 months of learning experience with distance perception. To control for this experience, it would have been necessary to raise infants in complete visual isolation, a procedure that was obviously out of the question—although, as you recall from Box 2.1, Dennis (1941) had few qualms about subjecting infants to an impoverished environment. Such an isolation procedure is feasible with animals, however, in part because the isolation does not have to be long—animals develop the ability to move through their environments very quickly, sometimes in a matter of minutes. So Gibson and Walk tested a variety of species from rats to kittens to lambs, isolating them from birth (i.e., no specific visual experiences) until they could move around competently and then testing them on the visual cliff. They discovered that depth perception, at least as measured in the cliff apparatus, is built into the visual system, at least for those species that rely heavily on vision.

Animal Rights

The use of animals in research is an emotional and controversial issue (not a new one, though—see Box 2.3). Animal rights activists have denounced the use of animals in studies ranging from medical research to cosmetics testing. The majority of animal activists confine their activities to sincere argument and nonviolent protest, and they work hard to live a life that is consistent with their moral stance (Herzog, 1993). In some cases, however, activism has led to animal laboratories being vandalized and animals released from labs. During the 1980s, for example, animal rights extremists vandalized approximately 100 research facilities housing animals (Adler, 1992). The problem was severe enough to produce federal legislation, the Animal Enterprise Protection Act of 1992, specifically outlawing such vandalism and setting stiff penalties, and the Animal Enterprise Terrorism Act of 2006, which took an even harder line. In recent years, an

alarming trend has been for some groups to target researchers directly, not just their labs. In the fall of 2008, for instance, two researchers at the University of California at Vera Cruz were the targets of firebombs (the car of one, the home of another).

BOX 2.3 ORIGINS—Antivivisection and the APA

Considering the high visibility of the animal research controversy, you might think that it is a fairly recent development. Not so; it has a long history, as documented nicely by the comparative psychologist and historian Donald Dewsbury (1990).

The term vivisection derives from the Latin vivus, or "alive," and refers to surgical procedures on live animals, historically done for scientific purposes. The antivivisection movement developed in 19th-century England, where activists' efforts contributed to the passage of England's Cruelty to Animals Act in 1876, an ethics code similar in spirit to modern APA guidelines for animals. The antivivisection movement quickly spread to the United States, where the American Antivivisection Society was founded in 1883 in Philadelphia. Antivivisectionists and animal researchers (including physiologists and early experimental psychologists) engaged in the same arguments that are heard today, with claims of unspeakable torture on the one side and justifications on scientific grounds on the other. That thoughtful scientists were torn by the issue of using animals in research is reflected in the experiences of Charles Darwin. An animal lover, surrounded by pets all his life, Darwin nonetheless argued for the importance of legitimate animal research, writing in 1871 that "it is justifiable for real investigations on physiology; but it is not for mere damnable and detestable curiosity" (quoted in Dewsbury, 1990, p. 316).

Within the field of psychology in the early years of the 20th century, one especially controversial series of animal studies concerned John B. Watson (again). In order to determine which senses were critical for maze learning, Watson conducted a series of studies in which he surgically eliminated one sense one at a time to examine the effects on rats in mazes (Watson, 1907). For instance, he learned that vision and smell did not affect the learning of a maze or the retention of an already-learned maze. Rats surgically blinded or with olfactory bulbs removed performed the same as unimpaired rats. The study caused an outcry when it was reported in the New York Times on December 30, 1906, and Watson was vilified in the antivivisectionist Journal of Zoophily, which also printed the cartoon shown in Figure 2.2 (from Dewsbury, 1990).

FIGURE 2.2 Antivivisectionist cartoon of Watson on the operating table. From Dewsbury (1990).

The APA established its first code for regulating animal research in the 1920s, well before creating the code for research with humans. A committee chaired by Robert Yerkes was formed in 1924, and the following year the APA adopted its recommendations. The committee proposed that laboratories create an open-door policy in which "any accredited member of a humane society [could] be permitted to visit a laboratory to

observe the care of animals and methods of experimentation" (Anderson, 1926, p. 125), that journals require authors to be clear about the use of humane procedures in their research, that psychologists defend the need for animal research, both in the classroom and publicly, and that the APA maintain a standing committee on "precautions in animal experimentation" (Anderson, 1926, p. 125).

What is the case against the use of animals as research subjects? Some argue that humans have no right to consider themselves superior to any other sentient species—that is, any species capable of experiencing pain (Singer, 1975). Sentient animals are said to have the same basic rights to privacy, autonomy, and freedom from harm as humans and therefore cannot be subjugated by humans in any way, including participation in any form of research. Others skeptical of animal research take a more moderate position, grounded in a Judeo-Christian theology. They argue that humans may have dominion over animals, but they also have a responsibility to protect them. This group recognizes the value of some research using animals, especially medical research, but rejects other types of experimentation on the grounds that researchers have inflicted needless pain and suffering when alternative approaches to the research would yield essentially the same conclusions. This argument has helped reduce unnecessary research on animals by the cosmetics industry, for instance, but it has been applied to research in psychology as well. Psychological research with animals has been described as needlessly repetitive and concerned with trivial problems that have no practical human benefit. Critics have suggested that instead of using animals in the laboratory, researchers could discover all they need to know about animal behavior by observing animals in their natural habitats, by substituting non-sentient for sentient animals, or by using computer simulations. How do research psychologists respond?

Using Animals in Psychological Research

Most psychologists simply do not agree that sentient animals have rights equal to those of humans. While granting that humans have an obligation to protect and care for nonhuman species, psychologists believe humans can be distinguished from nonhumans because of our degree of awareness, our ability to develop culture and understand history, and especially our ability to make moral judgments. Although animals are capable of complex cognition, they are "incapable of being moral subjects, of acting rightly or wrongly in the moral sense, of having, discharging, or breaching duties and obligations" (Feinberg, 1974, p. 46). Of course, differentiating between human and nonhuman species does not by itself allow the use of the latter by the former. Some psychologists (e.g., Ulrich, 1991) caution there have indeed been instances in which animals were not treated well by research psychologists and that some research has been needlessly repetitive. Most psychologists argue, however, that the use of animals in research does not constitute exploitation and that the net effect of such research is beneficial rather than costly for both humans and animals.

The most visible defender of animal research in psychology has been Neal Miller (1909–2002), a noted experimental psychologist. His research, on topics ranging from basic

processes in conditioning and motivation to the principles underlying biofeedback, earned him the APA's Distinguished Scientific Contributions Award in 1959 and its Distinguished Professional Contributions Award in 1983. In "The Value of Behavioral Research on Animals", Miller (1985) argued that (a) animal activists sometimes overstate the harm done to animals in psychological research, (b) animal research provides clear benefits for the well-being of humans, and (c) animal research benefits animals as well. Concerning harm, Miller cited a study by Coile and Miller (1984) that examined 5 years' worth of published research in APA journals, a total of 608 studies, and found no instances of the forms of abuse claimed by activists. Also, examining the abuse claims shows at least some of the alleged "abuse" may not be that at all, but merely seems to be because of the inflammatory language used. For instance, Coile and Miller cited several misleading statements from activist literature, including: "[The animals] are deprived of food and water to suffer and die slowly from hunger and thirst" (p. 700). This evidently refers to the common laboratory practice in conditioning experiments of depriving animals of food or water for 24 hours. Animals then placed in a conditioning procedure are motivated to work for the food or the water (e.g., solve a maze). Is this abuse? Perhaps not, considering that veterinarians recommend most pets be fed just once a day (Gallup & Suarez, 1985b). On the other hand, some researchers argue that 6 hours without food is sufficient to create an adequate level of hunger for research purposes.

Miller (1985) argued that situations involving harm to animals during research procedures are rare, used only when less painful alternatives cannot be used, and can be justified by the ultimate good that derives from the studies. This good applies to both humans and animals, and the bulk of his 1985 article was an attempt to document the kinds of good that derive from animal studies. First, he argued that while the long history of animal conditioning research has taught us much about general principles of learning, it also has had direct application to human problems. An early example of this was a device developed and tested by Mowrer and Mowrer (1938) for treating enuresis (excessive and uncontrolled bedwetting) that was based explicitly on the classical conditioning work involving Pavlov's dogs. Teaching machines and several forms of behavior therapy (e.g., systematic desensitization) are likewise grounded in conditioning principles originally observed in research with animals. More recently, animal research has directly influenced the development of behavioral medicine—the application of behavioral principles to traditional medical practice. Disorders ranging from headaches to hypertension to the disabilities following strokes can be treated with behavioral procedures such as biofeedback, and the essential principles of biofeedback were determined using animals as subjects.

Finally, Miller (1985) argued that animal research provides direct benefits to animals themselves. Medical research with animals has improved veterinary care dramatically (e.g., developing rabies vaccine), but behavioral research has also improved the welfare of various species. The study of animal behavior by research psychologists has led to improvements in the design of zoo environments, aided in nonchemical pest control, and discouraged coyote attacks on sheep by using taste avoidance conditioning as a substitute

for lethal control. Behavioral research can even help preserve endangered species. Miller used the example of imprinting, the tendency for young ducklings and other species to follow the first stimulus that moves (usually the mother). Research on imprinting led to the procedure of exposing newly hatched condors to a puppet resembling an adult condor rather than to a normal human caretaker, thereby facilitating the bonding process for the incubator-raised bird and ultimately enhancing the survival of this threatened species.

Another area of research involving animals, one that benefits both animals and humans, is anthrozoology—the study of human-animal interactions. The field is interdisciplinary, and includes behavioral psychologists, veterinarians, anthropologists, animal trainers, and philosophers. The topics they study include the use of pets in psychotherapy (decidedly mixed results—see Herzog, 2011), the effects of pets on the everyday lives of humans, and the training of both animals and humans to improve human–animal relationships. A good introduction, with a title that highlights the moral ambiguities of our complex relationships with animals is Some We Love, Some We Hate, Some We Eat: Why It's So Hard to Think Straight About Animals, by Hal Herzog (2010), a research psychologist who has become a leader in this emerging discipline.

One last point about using animals in psychological research is that most people seem to think animal research has value. Surveys of psychologists (Plous, 1996a) and psychology majors (Plous, 1996b), for instance, indicate that although they are ambivalent about research in which animals experience pain and/or must be put to death at the conclusion of the study, most psychologists and students of psychology believe that animal research in psychology is both justified and necessary. These views appear to be shared by students in general (Fulero & Kirkland, 1992; Gallup & Beckstead, 1988). Despite this general support, there are indications that animal research by psychologists is in decline. Gallup and Eddy (1990), for example, surveyed graduate programs and reported that 14.7% of them had dismantled their animal labs, mainly due to changing research interests and cost (and not because of pressure from protesters). Benedict and Stoloff (1991) found similar results among elite undergraduate colleges. The Plous surveys of psychologists and students just mentioned found general support for animal research, but his analysis also revealed stronger support among (a) older psychologists and (b) male psychologists and male psychology majors. This outcome suggests animal research among psychologists, as well as animal labs for undergraduate psychology majors (about 70% of whom are now female), may decline in the future. On the other hand, Hull (1996) surveyed 110 department chairs at schools with undergraduate psychology majors but without graduate programs and found some reason for optimism about animal research. She reported just a small drop in the use of animal labs for undergraduates over a 5-year period; 47% reported using animals (mainly rats) at the time of the survey, while 50% had used animals 5 years earlier. Hull's survey also revealed that departments using animals did not find APA and National Institutes of Health (NIH) guidelines difficult to follow and that the student response to the presence of an animal lab was mostly favorable.

The APA Code for Animal Research

Psychologists must follow federal, state, and local laws governing use of animals in research. The Animal Welfare Act (AWA) enacted in 1966 is the only federal law in the United States that regulates the treatment of animals used in research. Part of the AWA's mandate is that institutions where animal research is conducted should have an Institutional Animal Care and Use Committee (IACUC). Like an IRB, the IACUC is composed of faculty from several disciplines in addition to science, a veterinarian, and someone from outside the university.7 Often, the IACUC will use guidelines put forth in the Guide for the Care and Use of Laboratory Animals (National Research Council, 2011) in its evaluation of the ethical treatment of animals in research. In addition, psychologists rely on Standard 8.09 of the 2002 APA ethics code, which describes the ethical guidelines for animal care and use (www.apa.org/science/anguide.html). The APA guidelines for using animals deal with (a) the need to justify the study when the potential for harm to the animals exists; (b) the proper acquisition and care of animals, both during and after the study; and (c) the use of animals for educational rather than research purposes. The main theme of the code is balancing the scientific justification for a particular project with the potential for harm to the animals. Here are the highlights.

Justifying the Study

Just as the researcher studying humans must weigh the scientific value of the research against the degree of risk to the participants, the animal researcher must make the case that the "scientific purpose of the research [is] of sufficient potential significance as to outweigh any harm or distress to the animals used" (APA, 1985, p. 5). The scientific purpose of the study should fall within one of four categories. The research should "(a) increase knowledge of the processes underlying the evolution, development, maintenance, alteration, control, or biological significance of behavior, (b) determine the replicability and generality of prior research, (c) increase understanding of the species under study, or (d) provide results that benefit the health or welfare of humans or other animals" (www.apa.org/science/anguide.html).

The longest section of the guidelines identifies the range of procedures that can be used. In general, researchers are told that their requirement for a strong justification increases with the degree of discomfort to be experienced by the animals. In addition, they are told that appetitive procedures (i.e., use of positive reinforcement) should be substituted for aversive procedures as much as possible, that less stressful procedures should be preferred to more stressful ones, and that surgical procedures require special care and expertise. Researchers are also encouraged to try out painful procedures on themselves first, whenever feasible. Field research procedures should disturb animals living in their natural habitat as little as possible.

Caring for the Animals

The research supervisor must be an expert in the care of the species of animals to be used, must carefully train all those who will be in contact with the animals, and must be fully aware of federal regulations about animal care. To further ensure proper care, a veterinarian must check the facilities twice annually and be on call as a general consultant.

The animals should be acquired from legitimate suppliers or bred in the laboratory. If wild animals are studied in a laboratory, they must be trapped humanely.

Once an experiment is completed, alternatives to destroying the animals should be considered. However, euthanasia is sometimes necessary, "either as a requirement of the research, or because it constitutes the most humane form of disposition of an animal at the conclusion of the research" (APA, 1985, p. 8). In such cases, the process must be "accomplished in a humane manner, appropriate for the species, under anesthesia, or in such a way as to ensure immediate death, and in accordance with the procedures approved by the institutional animal care and use committee" (p. 8).

Using Animals for Educational Purposes

The guidelines are designed primarily to aid researchers who test animals, but animals are often used educationally to demonstrate specific behaviors, train students in animal research procedures, and give students firsthand experience in studying such well-known phenomena as classical and operant conditioning. Unlike the research situation, the educational use of animals does not result directly in new knowledge. Consequently, the educator is urged to use fewer rather than more animals to accomplish a given purpose and to consider a variety of alternative procedures. For example, instead of demonstrating the same principle (e.g., shaping) to an introductory psychology class with a new rat each semester, the instructor might do it once and make a video of the procedure for future classes.

Sometimes, computer simulations of phenomena can be substituted for live procedures; several reasonably accurate simulations of both classical and operant conditioning procedures exist. These simulations can be effective (and necessary in smaller schools that cannot keep up with federal regulations for the proper care of animals), but shaping a schematized rat to bar press is not the same as shaping a real rat. Students often experience a deep insight into the power of reinforcement contingencies when they witness the animals firsthand. Direct experiences with animals in undergraduate learning laboratories have motivated more than one student to become a research psychologist (Moses, 1991).

In summary, most psychologists defend the use of animals in behavioral research while recognizing the need to scrutinize closely the rationale for every animal study. Animal research has contributed greatly to our understanding of behavior and promises to help in the future search for solutions to AIDS, Alzheimer's disease, mental illness, and countless other human problems.

Scientific Fraud

There has been much discussion in recent years about fraud in science, with specific cases sparking debate about whether they represent the occasional bad apple or a broader, more systemic problem. Scientists in general and psychological scientists in particular are expected to be scrupulously honest in all of their scientific activities. Principle C (Integrity)

of the APA ethics code unambiguously states that psychologists "seek to promote accuracy, honesty, and truthfulness in the science, teaching, and practice of psychology" (APA, 2002, p. 1062). This last section of the chapter examines the issue of scientific fraud, aiming to shed some light on the varieties of scientific misbehavior, the reasons why such behavior occasionally occurs, and the ways in which fraud can be detected.

The American Heritage Dictionary (1992) defines fraud as "a deception deliberately practiced in order to secure unfair or unlawful gain" (p. 722). The two major types of serious misconduct in science are (1) plagiarism, deliberately taking the ideas of someone else and claiming them as one's own, and (2) falsifying data. In the APA ethics code, plagiarism is specifically condemned in Standard 8.11 – "Psychologists do not present portions of another's work or data as their own, even if the other work or data source is cited occasionally" (APA, 2002, p. 1070); data falsification receives similar treatment in Standard 8.10a – "Psychologists do not fabricate data" (p. 1070). Plagiarism is a problem that can occur in all disciplines, and you will find further discussion of it within the context of writing APA-style lab reports (See Appendix A). Being dishonest about data, on the other hand, is a problem that happens only in science; it will be our major focus.

Data Falsification

Data are the foundation on which the entire scientific enterprise is built. If there is a mortal sin in science, it is the failure to be scrupulously honest in collecting and managing data. The most egregious sin is data falsification, which occurs when a scientist fails to collect data and simply fabricates a data set, or collects some data, but either manufactures the rest or changes some of the existing data to produce a favorable outcome. Each of these forms of data falsification occurred in a notorious recent case.

Until his 2011 resignation, Diederick Stapel was dean of the School of Social and Behavioral Sciences at Tilburg University in the Netherlands, and he was a star, one of Europe's premier social psychologists. His research was published in numerous prominent journals and frequently cited. Many of the studies involved the phenomenon of "priming," in which subjects are presented with stimuli or put in environments that lead them to think unconsciously in some fashion and then behave in certain predictable ways. In one of Stapel's better known studies, for instance, subjects placed in a trash-filled environment seemed to show racist tendencies. Specifically, White subjects were asked to choose a seat in a row of six chairs and complete a questionnaire. Sitting in the seat at the end of the row was another apparent subject, who was either Black or White. According to Stapel's data, White subjects chose to sit further away from the Black person than from the White person. According to Stapel, the dirty environment primed racist tendencies in the White subjects, leading them to avoid sitting near the Black person.

This study was published in Science, among the most prestigious of all scientific journals and it quickly became widely known. The only problem was that Stapel never actually conducted the study and he made up all the data. He was smart enough, however, to create data that would not raise suspicions—the results were statistically significant, but

not so large as to be unbelievable. This was his typical way of committing scientific fraud—find an interesting topic, develop a simple yet creative procedure, and then manufacture results that were statistically significant but believable and simple to understand. He was even able to get away with his data falsification scheme when collaborating with other social psychologists and with graduate students. He would work closely with colleagues and students when designing a study, often helping them develop their own ideas into interesting research designs. Stapel would then tell them that the actual data collection would occur at nearby secondary schools rather than at the university, with the data collected by one of his many (fictitious, as it turned out) research friends. Stapel would then manufacture the data and give the results (not the raw data, but data coded in a way that it could be entered into a computer) to students or colleagues for analysis (Jump, 2011).

Stapel was eventually undone when he raised the suspicions of some of his more astute graduate students and colleagues. He never seemed to be able to produce raw data or the participants' completed questionnaires when asked for them (a common courtesy among scientists), his studies always seemed to work out as hypothesized (even the best researchers produce many studies that fail), and he eventually got sloppy in his data creation (one study included identical data in several places, an apparent cut-and-paste job). Stapel was investigated by his university, confessed, and resigned from Tilburg. No fewer than 55 of his scientific papers were retracted from journals (Bhattacharjee, 2013).

Explanations for why research fraud occurs range from individual (character weakness) to societal (a reflection of an alleged moral decay in modern society), with reasons relating to the academic reward system somewhere in the middle. Scientists who publish are promoted, tenured, win grants, and become influential in their fields. Sometimes, the pressure to "publish or perish" overwhelms the individual and leads the researcher (or the researcher's assistants) to cut some corners. The fraud might begin on a small scale—adding, subtracting, or altering a few pieces of data to achieve the desired outcome—but it may expand over time. Changing small amounts of data was the starting point for Stapel—long before he was manufacturing entire sets of data, he was changing individual data points to produce desired results (Crocker, 2011).

As for uncovering fraud based on falsified data, the traditional view is that it will be detected eventually because faked results won't be replicated (Hilgartner, 1990). That is, if a scientist produces a result with fraudulent data, the results won't represent some empirical truth. Hence, other scientists, intrigued or surprised by the new finding, will try to reproduce it in their own labs and may fail to do so. This will raise suspicions, and fraudulent findings eventually will be uncovered and discarded. Yet a failure to replicate is by no means a foolproof indication of fraud—results might not reproduce for several reasons. In addition, as occurred in the Stapel case, the clever fraudster can create data that will be entirely believable and may indeed replicate.

A failure to replicate may in some cases raise suspicion and subsequently lead to a request to see the raw data. However, failure to produce such data will generate even more suspicions. Scientists in psychology and other disciplines have a long history of willingness to share data and a refusal to do so would create concern about the new findings, as was the case in the Stapel fraud. Standard 8.14 of the ethics code makes it clear that data sharing is expected from researchers.

Psychological science is a collaborative activity (the concept of a "research team" will be elaborated in Chapter 3), and it is difficult to fool sharp-minded students and other research psychologists for too long. Although it took years in the Stapel case, graduate student suspicion, as we have seen, was the starting point for his downfall. Colleague suspicion was also the starting point in another well-known case of fraud that occurred in the 1980s, following a series of studies that apparently made a breakthrough in the treatment of hyperactivity in children with intellectual disabilitities. Stephen Breuning of the University of Pittsburgh produced data appearing to show that stimulant drugs could be more effective than antipsychotic drugs for treating the problem (Holden, 1987). However, a colleague suspected the data were falsified, a charge that was upheld after an investigation by the National Institute of Mental Health (NIMH), which had funded some of Breuning's research. In a plea bargain, Breuning pled guilty to two counts of submitting false data to NIMH; in exchange, NIMH dropped the charge that Breuning committed perjury during the investigation (Byrne, 1988).

It is worth mentioning that some commentators (e.g., Hilgartner, 1990) believe that while falsified data may go undetected for some time because they replicate "good" data or the data just seem believable, falsified data may not be detected for two other reasons as well. First, the sheer number of studies being published today makes it easier for a bad study to slip through the cracks, especially if it isn't reporting a notable discovery that attracts widespread attention. Second, the reward system in science is structured so that new discoveries pay off, but scientists who spend their time "merely" replicating other work aren't seen as creative. However, this latter issue is currently being addressed as researchers are now undergoing various systematic replication projects; this will be discussed in more detail in Chapters 3 and 4. A third issue is that researchers are "rewarded" by publishing counterintuitive or surprising findings, which may in turn garner media attention (Shea, 2011). While it is important that we "give psychology away" as George Miller recommended (see Chapter 1), it is imperative that the information shared with the larger community is based on sound, scientific practice.

What does all this mean for you as a student researcher? At the very least, it means you must be compulsive about data. Follow procedures scrupulously and never succumb to the temptation to manufacture or change even a single piece of data. Likewise, never discard data from a participant unless there are clear procedures for doing so and these procedures are specified before the experiment begins (e.g., the participant doesn't follow instructions, the experimenter doesn't administer the procedure correctly). Finally, keep the raw data or, at the very least, the data summary sheets. Your best protection

against a charge that your results seem unusual is your ability to produce the data on request. Being vigilant and truthful about your data will make you a better scientist and a better seeker of truth.

SELF TEST 2.3

Miller argued that animal rights activists exaggerate when making claims about animal research. What were his other two arguments for the value of animal research in psychology?

What does the APA recommend about the use of animals for educational purposes?

Which facts first alerted researchers to the possibility of fraud in Stapel's research?

The importance of being aware of the ethical implications of the research you're doing cannot be overstated. It is the reason for placing this chapter early in the text, and it won't be the last you'll hear of the topic. If you glance back at the table of contents, for instance, you will notice that each of the remaining chapters includes an Ethics Box that examines such topics as maintaining privacy in field research, recruiting participants, using surveys responsibly, and being an ethically competent experimenter. On the immediate horizon, however, is a chapter that considers the problem of how to begin developing ideas for research projects.

CHAPTER SUMMARY

Developing the APA Code of Ethics

In keeping with psychology's habit of relying on data-based principles, the APA developed its initial ethics code empirically, using a critical incidents procedure. The code for research using human participants was first published in 1953 and has been revised periodically since then, most recently in 2002. It consists of general principles guiding the behavior of psychologists (e.g., concern for others' welfare) and specific standards of behavior (e.g., maintaining the confidentiality of research participants), the violation of which can lead to censure.

Ethical Guidelines for Research with Humans

The APA code for research with humans provides guidance for the researcher in planning and carrying out the study. Planning includes doing a cost-benefit analysis that weighs the degree of risk imposed on participants against the scientific value of the research. The code also requires that subjects be given sufficient information to decide whether or not to participate (i.e., informed consent). Special care must be taken with children and with people who might feel coerced into participation (e.g., prisoners). Participants must be told that they are free to withdraw from the study without penalty, and they must be assured of the confidentiality of their responses. At the conclusion of their participation, they must receive a full debriefing. Institutional Review Boards (IRBs) are responsible for ensuring research studies with human subjects are conducted according to the ethics code and federal law. Certain forms of deception are acceptable in psychological research, but the researcher must convince an IRB the legitimate goals of the study can be met only through deception.

Ethical Guidelines for Research with Animals

APA guidelines for research with animal subjects concern the care and humane treatment of animals used for psychological research, provide guidance in choosing appropriate experimental procedures, and cover the use of animals both for research and for educational purposes. Although animal rights proponents have argued that animal research in psychology is inappropriate, most research psychologists argue that such research can benefit both humans and animals.

Scientific Fraud

Plagiarism (presenting the ideas of another as one's own) and data falsification (the manufacturing or altering of data) are the most serious forms of scientific fraud. Although data falsification is often discovered because of repeated failures to replicate unreliable findings, it may remain undetected because (a) the fraudulent findings are consistent with legitimate outcomes or (b) the sheer mass of published work precludes much replication. The academic reward system sometimes creates pressures that lead to scientific fraud.

CHAPTER REVIEW QUESTIONS

Distinguish between the general principles and the standards of the APA ethics code. Describe any three of the general principles, as they apply to research.

Describe the basic purpose of IRBs and the reasons research psychologists have criticized them.

What factors determine whether research proposals are exempt from IRB review, receive expedited review, or are subject to full review? How does the concept of risk relate to these judgments?

Distinguish between consent and assent and explain how both concepts are accomplished in research with children.

Describe the essential ingredients of an informed consent form to be used in research with adult participants.

Why is deception sometimes used in psychological research? How can the use of deception be reconciled with the concept of informed consent?

Describe the two main purposes of a debriefing session. When might a full debriefing be delayed until the experiment is completed?

Which ethical principles were violated in (a) the Willowbrook study, (b) the Tuskegee study, and (c) MK-ULTRA?

Use the Gibson visual cliff study to explain why psychologists sometimes use nonhuman species as research subjects.

Describe the arguments for and against the use of nonhuman species in psychological research.

Describe the kinds of research likely to be undertaken by anthrozoologists.

What are the essential features of the APA code for animal research?

What does the APA ethics code say about the use of animals for educational purposes?

Describe the ways in which data falsification is usually discovered. Why does this type of fraud occur?

APPLICATIONS EXERCISES

Exercise 2.1. Thinking Scientifically About Deception

From the standpoint of a research psychologist who is thinking scientifically, how would you design a study to evaluate the following claims that are sometimes made about the use of deception in research? That is, what kinds of empirical data would you like to have in order to judge the truth of the claims?

Deception should never be used in psychological research because once people have been deceived in a study, they will no longer trust any psychologist.

Researchers could avoid deception by instructing subjects to imagine they are in a deception study and then behave as they think a typical person would.

Psychologists are just fooling themselves; most participants see right through their deceptions and quickly understand the true purpose of a study.

Deception seldom works in research with university students, because they talk to each other about the studies in which they have participated and tell each other the "true" purpose of the studies.

Exercise 2.2. Recognizing Ethical Problems

Consider each of the following brief descriptions of actual research in social psychology. From the standpoint of the APA's code of ethics, which components could cause problems with an IRB? Explain how you might defend each study to an IRB.

The effect of crowding on stress was investigated in a public men's room. A member of the research team followed a subject into the bathroom and occupied either the urinal directly adjacent to the subject's or the next one down the line. Subjects were unaware they were participating in a study. On the assumption that increased stress would affect urination, the amount of time it took for the subject to begin to urinate and the total time spent urinating were recorded by another researcher hidden in one of the stalls. As predicted, subjects' urination was more disrupted when the immediately adjacent urinal was occupied (Middlemist, Knowles, & Matter, 1976).

In a field experiment, a woman (who was actually part of the experiment) stood by her car on the side of a road. The car had a flat tire. To determine if modeling would affect the helping behavior of passing motorists, on some trials another woman with a flat tire was helped by a stopped motorist (all part of the staged event) about a quarter-mile before the place where the woman waited for help. As expected, motorists were more likely to stop and help if they had just witnessed another person helping (Bryan & Test, 1967).

In the wake of the Watergate scandal, researchers wished to determine if average people could be induced to commit a crime, especially if they thought an arm of government would give them immunity from prosecution. Subjects were recruited by the experimenter, posing as a private investigator, and asked to be part of a break-in at a local advertising agency said to be involved in tax fraud. Some subjects were told that the IRS was organizing the break-in and promised immunity from prosecution; others weren't promised immunity. A third group was told a competing advertising agency was leading the break-in, and a fourth group was not told who was behind the crime. The prediction that people would be most willing to participate for a government agency that promised

immunity was confirmed; the experiment ended when participants either agreed or disagreed. No break-in actually occurred (West, Gunn, & Chernicky, 1975).

Exercise 2.3. Replicating Milgram

Describe what changes you think could be made to Milgram's basic obedience study in order to get it approved by an IRB today. Then track down Burger's description of his replication (Burger, 2009) and describe exactly what he did. On the basis of his study, do you think it is safe to conclude that Milgram's studies have been replicated and people are just as obedient today as they were in the 1960s?

Exercise 2.4. Decisions about Animal Research

Decisions about Animal Research The following exercise is based on a study by Galvin and Herzog (1992) and is used with the permission of Hal Herzog. The idea is for you to play the role of a member of an IACUC (Institutional Animal Care and Use Committee) and make decisions about whether the following studies ought to gain IACUC approval (quoting from Galvin & Herzog, p. 265):

Mice. A neurobiologist proposes to amputate the forelimbs of newborn mice to study the relative influence of heredity and environment on the development of motor patterns (grooming).

Rats. A psychologist seeks permission to conduct a classroom learning demonstration. Rats are to be deprived of food for 23 hours and taught to press a lever for food reinforcements.

Monkeys. Tissue from monkey fetuses will be implanted into the brains of adult rhesus monkeys to explore the feasibility of neural transplantation as a treatment for Alzheimer's disease.

Dogs. Stray dogs awaiting euthanasia in an animal shelter are to be used to teach surgical techniques to veterinary students.

Bears. Wild grizzly bears will be anesthetized. Collars containing radio telemetry devices will be attached to their necks for a study of their social and territorial behavior patterns.

For each of these studies, do a cost-benefit analysis, indicate whether you would approve the study, and explain the reasons why or why not. In terms of the ethics code, indicate whether some changes in procedure might switch your decision from "reject" to "approve."

ANSWERS TO SELF TESTS

✓ 2.1

The Hobbs committee used the procedure to collect examples of perceived ethical violations among psychologists.

They believed the infant has a strong constitution and would not be harmed; they also believed that the contribution to science outweighed any minor discomfort they would cause.

Respect for Persons, Beneficence, Justice

It means researchers must always weigh the benefits of their research against the potential harm to subjects, in order to achieve the greatest good.

✓ 2.2

Expedited review.

The potential subject is given enough information about the study to make a reasoned decision about whether or not to participate.

His procedure violated the “quit any time” proviso.

✓ 2.3

It benefits the well-being of humans; it also benefits animals (e.g., zoos).

Use live animals as sparingly as possible.

A failure to produce his raw data when asked and reproducing identical results across studies.

Notes

1 Humans are animals, too, of course. When we use the term animal research, we are referring to research with nonhuman animals.

2 The APA has established procedures for evaluating claims of ethical misconduct and for punishing those found guilty of misconduct. There is even a link allowing psychologists to report “critical incidents.” For more information, visit http://www.apa.org/ethics/code/.

3 Another useful source of information about the ethical treatment of human research participants is the Office for Human Research Protections in the U. S. Department of Health and Human Services. Its website is www.hhs.gov/ohrp/.

4 The question of what to call those who participate in psychological research has changed over the years. In the early 20th century, during the era of introspection (Chapter 1), participants were often called observers because their task was to observe what was going on in their minds during some task and then give an introspective report of it. As introspection went out of vogue, participants began to be called subjects. Starting with the fourth edition of its publication manual, in 1994, however, the APA mandated a change in this usage for articles published in APA journals. At least with regard to most humans (nonhuman animals and preverbal infants were still to be referred to as subjects), APA required writers to use research participant or participant instead of subject, apparently on the grounds that the latter term was somehow biased and dehumanizing. This change was widely criticized (e.g., Roediger, 2004) on the grounds that the term subject does not necessarily demean anyone, is more efficient linguistically (two as opposed to four syllables), and reflects historical continuity. In its most recent publication manual, APA (2010) has backed off and recognized the historical argument, noting that “for more than 100 years the term subjects has been used as a general starting point for describing a sample, and its use is appropriate (p. 73, italics in the original).

5 It is common practice for IRBs to include a research psychologist on the grounds that a substantial number of proposals come from psychology departments (Cook & Hoas, 2011).

6 For years, the Milgram study has been considered the “experiment that could never be replicated” because of the ethical issues involved. Recently, however, such a replication did occur at Santa Clara University (Burger, 2009), although substantial modifications were made to Milgram’s original procedure (e.g., not insisting that subjects continue, stopping the experiment earlier than Milgram did, screening subjects carefully). Despite the

methodological changes designed to placate his university IRB, the study found levels of obedience similar to those observed by Milgram.

7 As are IRBs, animal use committees have been controversial. One study found, for instance, that the same proposals given to different IACUCs yielded inconsistent levels of approval (Plous & Herzog, 2001).