9
Social Psychology as a Science

When I was in college, I first got interested in social psychology because it dealt with some of the most exciting aspects of being human: love, hate, prejudice, aggression, altruism, social influence, conformity, and the like. At that time, I didn't care a great deal about how this impressive body of knowledge came into existence. I simply wanted to know what was known. It wasn't until I entered graduate school that it suddenly dawned on me that I could be more than a consumer of this knowledge—I could become a producer, as well. And a whole new world opened up for me—the world of scientific social psychology. I learned how to ask important questions and do the experiments to find the answers to those questions—contributing, in my own small way, to the body of knowledge that I had read about as a student. And I have been passionately involved in that activity ever since.

Reading this chapter is not going to make you into a scientist. My intention for you is a bit less ambitious but no less important. This chapter is aimed at helping to improve your ability to think scientifically about things that are happening in your own social world. I have always found this a useful thing to be able to do. But, occasionally, it can be disillusioning, as well. Let me give you one example of what I mean by that statement. Several years ago, I picked up a copy of *The New Yorker* magazine, in which I read an excellent, highly informative essay by James Kunen[1] about college-level educational programs in our prisons. Kunen wrote enthusiastically about their effectiveness.

He then went on to decry the fact that a generally punitive congressional majority was eliminating these programs after characterizing them as wasteful and as tending to coddle criminals.

Kunen's essay contains a few vivid case histories of convicts who, while in prison, completed the college program and went on to lead productive lives after being released. The case histories are heartwarming. But, as a scientist, I wanted to know if there were any systematic data that I might use to evaluate the overall effectiveness of the program. Well, yes. Kunen reported one study published in 1991 by the New York State Department of Correctional Services, which found that 4 years after their release from prison, the recidivism rate of male inmates who had completed 1 or more years of higher education in prison was 20 percent lower than the average for all male inmates.

That sounds pretty impressive, right? Let's take a closer look. As scientists we need to ask one basic and vital question: Prior to participating in the program, were the prisoners who signed up for the program similar to those who didn't sign up? Might it not be the case that the prisoners who signed up for the program and completed a year of it were different *to begin with* (say, in motivation, ability, intelligence, prior education, mental health, or what have you) from those who did not sign up? I hasten to add that this is not simply nit-picking; if they were different at the outset from the general run of prisoners, then it is likely (or, at least, possible) that they would have had a lower rate of recidivism even without having taken the course of study. If that were the case, then it wasn't the program that caused the lower recidivism.

While I was reading Kunin's article, the liberal/humanist in me wanted to get excited by the results of this study; it would be terrific to have convincing data proving that educating prisoners pays off. But alas, the scientist in me took over and was skeptical. Thus, looking at the social world through the eyes of a scientist can be disillusioning. But it also gives us the ability to separate the wheat from the chaff so that, as concerned citizens, we can demand that innovative programs be properly evaluated. In that way, we can determine, with some degree of clarity, which of thousands of possible programs are worthy of our time, effort, and money. And the truth is that, in most cases, it is not difficult to do the experiment properly—as you will see.

What Is the Scientific Method?

The scientific method—regardless of whether it is being applied in physics, chemistry, biology, or social psychology—is the best way we humans have of satisfying our hunger for knowledge and understanding. More specifically, we use the scientific method in an attempt to uncover lawful relationships among things—whether the things are chemicals, planets, or the antecedents of human prejudice or love. The first step in the scientific process is observation. In physics, a simple observation might go something like this: If there is a rubber ball in my granddaughter's wagon and she pulls the wagon forward, the ball seems to roll to the back of the wagon. (It doesn't actually roll backward; it only seems that way.) When she stops the wagon abruptly, the ball rushes to the front of the wagon. In social psychology, a simple observation might go something like this: When I am waiting on tables, if I happen to be in a good mood and smile a lot at my customers, my tips seem to be a bit larger than when I am in a foul mood and smile less frequently.

The next step is to make a guess as to why that happens; this guess is our taking a stab at uncovering the "lawful relationship" we mentioned above. The third step is to frame that guess as a testable hypothesis. The final step is to design an experiment (or a series of experiments) that will either confirm or disconfirm the hypothesis. If a series of well-designed, well-executed experiments fails to confirm that hypothesis, we give it up. As my favorite physicist, Richard Feynman,[2] once put it, "It doesn't matter how beautiful the guess is or how smart the guesser is, or how famous the guesser is; if the experiment disagrees with the guess, then the guess is wrong. That's all there is to it." In my own opinion, this is both the essence of science and its beauty. There are no sacred truths in science.

Science and Art In my opinion, there is plenty of room for art in our science. I believe that the two processes—art and science—are different, but related. Pavel Semonov, a distinguished Russian psychologist, did a pretty good job of defining the difference. According to Semonov,[3] as scientists, we look closely at our environment and try to organize the unknown in a sensible and meaningful way. As artists, we reorganize the known environment to create something entirely new. To this observation, I would add that the requirements of a good

experiment frequently necessitate a combination of skills from both of these domains. In a very real sense, as experimenters, we use artistry to enrich our science. I believe this to be particularly true of experiments in social psychology.

Why is this blending of art and science especially true of social psychology? The full answer to this question will emerge as this chapter unfolds. For now, let me simply state that, in social psychology, we are not studying the behavior of chemicals in a beaker or of rubber balls in wagons; we are investigating the behavior of intelligent, curious, sophisticated adults who have been living in a social world for their entire lives. It goes without saying that, like the experimenters who are studying them, the people who serve as participants in our experiments have developed their own ideas and theories about what causes their feelings and behavior, as well as the feelings and behavior of the people around them. This is not the case when you are performing experiments with chemicals, with laboratory animals, or even with humans in nonsocial situations.

The fact that we are dealing with socially sophisticated human beings is part of what makes social psychology so fascinating as a topic of experimental investigation. At the same time, this situation also demands a great deal of art if the experimenter stands a chance of generating valid and reliable findings. In this chapter, I will try to communicate exactly how this happens.

From Speculation to Experimentation

In Chapter 8, we described a confusing phenomenon that we had stumbled upon several years ago: While John F. Kennedy was president, his personal popularity increased immediately after he committed a stupendously costly blunder. Specifically, after Kennedy's tragic miscalculation known as the Bay of Pigs fiasco, a Gallup poll showed that people liked him better than they had prior to that incident. Like most people, I was dumbfounded by this event. How could we like a guy better after he screwed up so badly? As a scientist, I speculated about what could have caused that shift. My guess was that, because Kennedy previously had been perceived as such a nearly perfect person, committing a blunder might have made him seem more human, thus allowing ordinary people to feel closer to him. An interesting speculation, but was it true?

Because many things were happening at the time of the Bay of Pigs fiasco, it was impossible to be sure whether this speculation was accurate. How might we have tried to find out? Well, we might have simply asked people why they liked Kennedy more now than they did the prior week. That sounds simple enough. Unfortunately, it is not that easy. Over the years, we have learned that people are often unaware of why they act in certain ways or change their beliefs in one direction or another; so, in a complex situation, simply asking people to explain their behavior will usually not yield reliable results.[4] This is precisely why social psychologists perform experiments. But how could we conduct an experiment on John F. Kennedy's popularity? We couldn't. In a case like this, we would try to conduct an experiment on the underlying phenomenon, not on the specific instantiation of that phenomenon. And, indeed, it was really the underlying phenomenon—not the specific event—that held our interest: Does committing a blunder increase the popularity of a nearly perfect person?

To answer this more general question, it was necessary to go beyond the event that led to our speculations. My colleagues and I[5] had to design an experiment that allowed us to control for extraneous variables and test the effects of a blunder on attraction in a less complex situation—one in which we could control the exact nature of the blunder, as well as the kind of person who committed it. And in that simple situation we found, as predicted, that "nearly perfect" people become *more* attractive after they commit a blunder, while "rather ordinary" people become *less* attractive after committing the identical blunder. (I have described the details of this experiment in Chapter 8.)

Designing an Experiment As suggested above, in striving for control, the experimenter must bring his or her ideas out of the helter-skelter of the real world and into the rather sterile confines of the laboratory. This typically entails concocting a situation bearing little resemblance to the real-world situation from which the idea originated. In fact, a frequent criticism is that laboratory experiments are unrealistic, contrived imitations of human interaction that don't reflect the real world at all. How accurate is this criticism?

Perhaps the best way to answer this question is to examine one laboratory experiment in great detail, considering its advantages and disadvantages, as well as an alternative, more realistic approach that

might have been used to study the same issue. The initiation experiment I performed in collaboration with Judson Mills[6] suits our purpose admirably—because it contains many of the advantages and disadvantages of the laboratory. The reader may recall that Mills and I speculated that people might come to like things for which they have suffered. We then designed and conducted a laboratory experiment in which we showed that people who expended great effort (by undergoing a severe initiation) to gain membership in a group liked the group more than did people who became members with little or no effort. Here's how the experiment was performed.

> Sixty-three college women who initially volunteered to engage in several discussions on the psychology of sex were participants of the study. Each student was tested individually. At the beginning of the study, I explained that I was studying the "dynamics of the group-discussion process." I said the actual topic of the discussion was not important to me, but because most people are interested in sex, I selected that topic to be certain of having plenty of participants. I also explained that I had encountered a major drawback in choosing sex as the topic: Specifically, because of shyness, many people found it difficult to discuss sex in a group setting. Because any impediment to the flow of the discussion could seriously invalidate the results, I needed to know if the participants felt any hesitancy to enter a discussion about sex. When the participants heard this, each and every one indicated she would have no difficulty. These elaborate instructions were used to set the stage for the important event to follow. The reader should note how the experimenter's statements tend to make the following material believable.
>
> Up to this point, the instructions had been the same for all participants. Now it was time to give each of the people in the various experimental conditions a different experience—an experience the experimenters believed would make a difference.
>
> Participants were randomly assigned in advance to one of three conditions: (1) One third of them would go through a severe initiation, (2) one third would go through a mild initiation, and (3) one third would not go through any initiation at all. For the no-initiation condition, participants were simply told they could now join the discussion group. For the severe- and mild-

initiation conditions, however, I told each participant that, because it was necessary to be positive she could discuss sex openly, I had developed a screening device—a test for embarrassment—that I then asked her to take. This test constituted the initiation. For the severe-initiation condition, the test was highly embarrassing. It required the participant to recite a list of 12 obscene words and 2 detailed descriptions of sexual activity taken from contemporary novels. The mild-initiation participants had to recite only a list of words related to sex that were not obscene.

The three conditions to which participants were assigned constituted the **independent variable** in this study. Briefly, the investigator's goal in designing and conducting an experiment is to determine if what happens to participants has an effect on how they respond. *Our* goal was to determine if severity of initiation—the independent variable—*caused* systematic differences in participants' behavior. Would participants who experienced a severe initiation act differently than those who experienced a mild initiation or no initiation at all?

But act differently in what way? After the initiation, each participant was allowed to eavesdrop on a discussion being conducted by members of the group she had just joined. To control the content of this material, a tape recording was used; but the participants were led to believe it was a live discussion. Thus, all participants—regardless of whether they had gone through a severe initiation, a mild initiation, or no initiation—listened to the same group discussion. The group discussion was as dull and as boring as possible; it involved a halting, inarticulate analysis of the secondary sex characteristics of lower animals—changes in plumage among birds, intricacies of the mating dance of certain spiders, and the like. The tape contained long pauses, a great deal of hemming and hawing, interruptions, incomplete sentences, and so on, all designed to make it boring.

At the end of the discussion, I returned with a set of rating scales and asked the participant to rate how interesting and worthwhile the discussion had been. This is called the **dependent variable** because, quite literally, the response is assumed to be "dependent" on the particular experimental conditions the participant had been assigned to. The dependent variable is what the experimenter measures to assess the effects of the

> independent variable. In short, if the independent variable is the *cause,* then the dependent variable is the *effect.*
>
> The results supported the hypothesis: Women who went through a mild initiation or no initiation at all saw the group discussion as relatively dull. But those who suffered in order to be admitted to the group thought it was really exciting. Remember, all the students were rating *exactly the same discussion.*

Designing and conducting this experiment was a laborious process. Mills and I spent hundreds of hours planning it, creating a credible situation, writing a script for the tape recording of the group discussion, rehearsing the actors who played the roles of group members, constructing the initiation procedures and the measuring instruments, recruiting volunteers to serve as participants, pilot-testing the procedure, running the participants through the experiment, and explaining the true purpose of the experiment to each participant (the reason for the deception, what it all meant, and so forth). What we found was that people who go through a severe initiation to join a group like that group a great deal more than people who go through a mild initiation or no initiation at all.

Surely there must be a simpler way! There is. The reader may have noticed a vague resemblance between the procedure used by Mills and me and other initiations, such as those used by primitive tribes and those used by some college fraternities and other exclusive clubs or organizations. Why, then, didn't we take advantage of the real-life situation, which is not only easier to study but also far more dramatic and realistic? Let's look at the advantages. Real-life initiations would be more severe (i.e., they would have more impact on the members); we would not have had to go to such lengths to design a group setting the participants would find convincing; the social interactions would involve real people rather than mere voices from a tape recording; we would have eliminated the ethical problem created by the use of deception and the use of a difficult and unpleasant experience in the name of science; and, finally, it could all have been accomplished in a fraction of the time the experiment consumed.

Thus, when we take a superficial look at the advantages of a natural situation, it appears that Mills and I would have had a much simpler job if we had studied existing fraternities. Here is how we

might have done it. We could have rated each group's initiation for severity and interviewed the members later to determine how much they liked their group. If the members who had undergone a severe initiation liked their fraternities more than the mild- or no-initiation fraternity members, the hypothesis would be supported. Or would it? Let's take a closer look at why people bother to do experiments.

If people were asked to name the most important characteristic of a laboratory experiment, the great majority would say "control." And this *is* a major advantage. Experiments have the advantage of controlling the environment and the variables so that the effects of each variable can be precisely studied. By taking our hypothesis to the laboratory, Mills and I eliminated a lot of the extraneous variation that exists in the real world. The severe initiations were all equal in intensity; this condition would have been difficult to match if we had used several severe-initiation fraternities. Further, the group discussion was identical for all participants; in the real world, however, fraternity members would have been rating fraternities that were, in fact, different from one another. Assuming we had been able to find a difference between the severe-initiation and mild-initiation fraternities, how would we have known whether this was a function of the initiation rather than of the differential likableness that already existed in the fraternity members themselves? In the experiment, the only difference was the severity of the initiation, so we know that any difference was due to that procedure.

The Importance of Random Assignment

Control *is* an important aspect of the laboratory experiment, but it is not the major advantage. A still more important advantage is that participants can be randomly assigned to the different experimental conditions. This means each participant has an equal chance to be in any condition in the study. Indeed, the **random assignment** of participants to conditions is the crucial difference between the experimental method and nonexperimental approaches. And the great advantage of the random assignment of people to conditions is this: Any variables not thoroughly controlled are, in theory, distributed randomly across the conditions. This means it is extremely unlikely that such variables would affect results in a systematic fashion.

An example might help to clarify this point: Suppose you are a scientist and you have the hypothesis that marrying intelligent women makes men happy. How do you test this hypothesis? Let us say you proceed to find 1,000 men who are married to intelligent women and 1,000 men who are married to not-so-intelligent women, and you give them all a "happiness" questionnaire. Lo and behold, you find that the men married to intelligent women *are* happier than the men married to less intelligent women. Does this mean that being married to an intelligent woman makes a man happy? No. Perhaps happy men are sweeter, more good-humored, and easier to get along with, and that, consequently, intelligent women seek out these men and marry them. So it may be that being happy *causes* men to marry intelligent women. The problem doesn't end there. It is also possible that there is some third factor that causes *both* happiness *and* being married to an intelligent woman. One such factor could be money: It is conceivable that being rich helps make men happy and that their being rich is what attracts the intelligent women. So it is possible that neither causal sequence is true. It is possible that happiness does not cause men to marry intelligent women and that intelligent women do not cause men to be happy.

The problem is even more complicated because we usually have no idea what these third factors might be. In the case of the happiness study, it could be wealth; it could also be that a mature personality causes men to be happy and also attracts intelligent women; it could be social grace, athletic ability, power, popularity, using the right toothpaste, being a snappy dresser, or any of a thousand qualities the poor researcher does not know about and could not possibly account for. But if the researcher performs an experiment, he or she can randomly assign participants to various experimental conditions. Although this procedure does not eliminate differences due to any of these variables (money, social grace, athletic ability, and the like), it neutralizes them by distributing these characteristics randomly across various experimental conditions. That is, if participants are randomly assigned to experimental conditions, there will be approximately as many rich men in one condition as in the others, as many socially adept men in one condition as in the others, and as many athletes in one condition as in the others. Thus, if we do find a difference between conditions, it is unlikely that this would be due to individual differences in any single characteristic because all of these

characteristics had an equal (or nearly equal) distribution across all of the conditions.

Admittedly, the particular example of intelligent women and their happy husbands does not easily lend itself to the confines of the experimental laboratory. But let us fantasize about how we would do it if we could. Ideally, we would take 50 men and randomly assign 25 to intelligent wives and 25 to less intelligent wives. A few months later, we could come back and administer the happiness questionnaire. If the men assigned to the intelligent wives are happier than the men assigned to the less intelligent wives, we would know what caused their happiness—we did! In short, their happiness couldn't easily be attributed to social grace, or handsomeness, or money, or power; these were randomly distributed among the experimental conditions. It almost certainly was caused by their wives' characteristics.

To repeat, this example is pure fantasy; even social psychologists must stop short of arranging marriages for scientific purposes. But this does not mean we cannot test important, meaningful, relevant events under controlled laboratory conditions. This book is loaded with such examples. Let's look at one of these examples as a way of clarifying the advantages of the experimental method. In Chapter 6, I reported a correlation between the amount of time children spend watching violence on television and their tendency to choose aggressive solutions to their problems.

Does this mean watching aggression on television causes youngsters to become aggressive? Not necessarily. It might. But it might also mean that aggressive youngsters simply like to watch aggression, and they would be just as aggressive if they watched *Sesame Street* all day long. But then, as we saw, some experimenters came along and proved that watching violence increases violence.[7] How? By randomly assigning some children to a situation in which they watched a video of an episode of a violent TV series—an episode in which people beat, kill, rape, bite, and slug each other for 25 minutes. As a control, the experimenters randomly assigned some other children to a situation in which they watched an athletic event for the same length of time. The crucial point: Each child stood *an equal chance* of being selected to watch the violent video as the nonviolent video; therefore, any differences in character structure among the children in this experiment were neutralized across the two experimental conditions. Thus, the finding that youngsters who watched the violent

video showed more aggression afterward than those who watched the athletic event suggests quite strongly that watching violence can lead to violence.

You may recall that this was precisely the problem with the evaluation of the prison college program that we described at the beginning of this chapter: The prisoners who volunteered for the program were probably different in many ways from those who did not volunteer. So it was misleading to compare their recidivism rate with that of the nonvolunteers. Such a comparison would stack the deck, making the program appear more effective than it actually was. How do you solve that problem? One way would be to attract twice as many volunteers for the program as you can handle. Then you can randomly select half of the volunteers for the program and place the other half in the control condition. If the selection is truly random, comparing the recidivism rate of the two groups would give you meaningful data.

Let us return to the initiation experiment. If we conducted a survey and found that members of severe-initiation fraternities find each other more attractive than do members of mild-initiation fraternities, then we would have evidence that severity of initiation and liking for other members of the fraternity are *positively correlated.* This means that the more severe the initiation, the more a member will like his fraternity brothers. No matter how highly correlated the two variables are, however, we cannot conclude, from our survey data alone, that severe initiations *cause* liking for the group. All we can conclude from such a survey is that these two factors are associated with each other.

It is possible that the positive correlation between severe initiation and liking for other members of a fraternity exists not because severe initiations cause members to like their groups more, but for just the opposite reason. It could be that the high attractiveness of the group causes severe initiations. If group members see themselves as highly desirable, they may try to keep the situation that way by maintaining an elite group. Thus, they may require a severe initiation to discourage people from joining unless those people have a strong desire to do so. From our survey data alone, we cannot conclude that this explanation is false and that severe initiations really do lead to liking. The data give us no basis for making this choice because they tell us nothing about cause and effect. Moreover, as we have seen in

our previous example, there could be a third variable that causes both severe initiations and liking. Who would like to give and receive a severe initiation? Why, people with strong sadomasochistic tendencies, of course. Such people may like one another not because of the initiation but because "birds of a feather" tend to like one another. Although this may sound like an outlandish explanation, it is certainly possible. What is more distressing for the researcher are the countless other explanations he or she can't even think of. The experimental method, based as it is on the technique of random assignment to experimental conditions, eliminates all of these in one fell swoop. The sadomasochists in the experiment have just as much chance of being assigned to the no-initiation condition as to the severe-initiation condition. In the real-world study, alas, almost all of them would assign themselves to the severe-initiation condition, thus making the results uninterpretable.

The Challenge of Experimentation in Social Psychology

Control Versus Impact All is not so sunny in the world of experimentation. There are some very real problems connected with doing experiments. I mentioned that control is one of the major advantages of the experiment, yet it is impossible to exercise complete control over the environment of human participants. One of the reasons many psychologists work with rats rather than people is that researchers are able to control almost everything that happens to their participants from the time of their birth until the experiment ends—climate, diet, exercise, degree of exposure to playmates, absence of traumatic experiences, and so on. Social psychologists do not keep human participants in cages to control their experiences. Although this makes for a happier world for the participants, it also makes for a slightly sloppy science.

Control is further limited by the fact that individuals differ from one another in countless subtle ways. Social psychologists try to make statements about what *people* do. By this we mean, of course, what most people do most of the time under a given set of conditions. To the extent that unmeasured individual differences are present in our results, our conclusions may not be precise for all people.

Differences in attitudes, values, abilities, personality characteristics, and recent experiences can affect the way people respond in an experiment. Thus, even with our ability to control the experimental situation itself, the same situation may not affect each person in exactly the same way.

Furthermore, when we do succeed in controlling the experimental setting so that it is exactly the same for every person, we run the real risk of making the situation so sterile that the participant is inclined not to take it seriously. The word *sterile* has at least two meanings: (1) germ-free, and (2) ineffective or barren. The experimenter should strive to make the experimental situation as "germ-free" as possible without making it barren or unlifelike for the participant. If participants do not find the events of an experiment interesting and absorbing, chances are their reactions will not be spontaneous and our results, therefore, will have little meaning. Thus, in addition to control, an experiment must have an impact on the participants. They must take the experiment seriously and become involved in it, lest it not affect their behavior in a meaningful way. The difficulty for social psychologists is that these two crucial factors, impact and control, often work in opposite ways: As one increases, the other tends to decrease. The dilemma facing experimenters is how to maximize the impact on the participants without sacrificing control over the situation. Resolving this dilemma requires considerable creativity and ingenuity in the design and construction of experimental situations. This leads us to the problem of realism.

Realism Early in this chapter, I mentioned that a frequent criticism of laboratory experiments is that they are artificial and contrived imitations of the world—that they aren't "real." What do we mean by *real?* Several years ago, in writing a treatise about the experimental method, Merrill Carlsmith and I[8] tried to pinpoint the definition of *real.* We reasoned that an experiment can be realistic in two separate ways: If an experiment has an impact on the participants, forces them to take the matter seriously, and involves them in the procedures, we can say it has achieved **experimental realism**. Quite apart from this is the question of how similar the laboratory experiment is to the events that frequently happen to people in the outside world. Carlsmith and I called this **mundane realism**. Often, confusion between experimental realism and mundane realism is responsible for the crit-

icism that experiments are artificial and worthless because they don't reflect the real world.

The difference between the two realisms can best be illustrated by providing you with an example of a study high in experimental realism but low in mundane realism. Recall the experiment by Stanley Milgram,[9] discussed in Chapter 2, in which each participant was asked to deliver shocks of increasing intensity to another person who was supposedly wired to an electrical apparatus in an adjoining room. Now, honestly, how many times in everyday life are we asked to deliver electric shocks to people? It's unrealistic—but only in the mundane sense. Did the procedure have experimental realism—that is, were the participants wrapped up in it, did they take it seriously, did it have an impact on them, was it part of their real world at that moment? Or were they merely playacting, not taking it seriously, going through the motions, ho-humming it? Milgram reports that his participants experienced a great deal of tension and discomfort. But I'll let Milgram describe, in his own words, what a typical participant looked like.

> I observed a mature and initially poised businessman enter the laboratory smiling and confident. Within 20 minutes he was reduced to a twitching, stuttering wreck, who was rapidly approaching a point of nervous collapse. He constantly pulled on his earlobe, and twisted his hands. At one point he pushed his fist onto his forehead and muttered: "Oh God, let's stop it." And yet he continued to respond to every word of the experimenter, and obeyed to the end.[9]

This hardly seems like the behavior of a person in an unrealistic situation. The things happening to Milgram's participants were *real*—even though they didn't happen to them in their everyday experience. Accordingly, it would seem safe to conclude that the results of this experiment are a reasonably accurate indication of the way people would react if a similar set of events *did* occur in the real world.

Deception The importance of experimental realism can hardly be overemphasized. The best way to achieve this essential quality is to design a setting that will be absorbing and interesting to the participants. At the same time, it is frequently necessary to disguise the true purpose of the study. Why the need for disguise?

Early in this chapter, I mentioned that just about everybody is an amateur social psychologist in the sense that we all live in a social world and are constantly forming hypotheses about things that happen to us in our social world. This includes the individuals who serve as participants in our experiments. Because they are always trying to figure things out, if they knew what we were trying to get at, they might be apt to behave in a manner consistent with their own hypotheses—instead of behaving in a way that is natural and usual for them. For this reason, we try to conceal the true nature of the experiment from the participants. Because we are almost always dealing with very intelligent adults, this is not an easy task; but it is an absolute requirement in most experiments if we are to stand a chance of obtaining valid and reliable data.

This requirement puts the social psychologist in the position of a film director who's setting the stage for action but not telling the actor what the play is all about. Such settings are called **cover stories** and are designed to increase experimental realism by producing a situation in which the participant can act naturally, without being inhibited by knowing just which aspect of behavior is being studied. For example, in the Aronson-Mills initiation study, participants were told they were taking a test for embarrassment to screen them for membership in a group that would be discussing the psychology of sex; this was the cover story. It was pure deception. In reality, they were being subjected to an initiation to see what effect, if any, this would have on their liking for the group. If the participants had been aware of the true purpose of the study before their participation, the results would have been totally meaningless. Researchers who have studied this issue have shown that, if participants know the true purpose of an experiment, they do not behave naturally but either try to perform in a way that puts themselves in a good light or try to "help out" the experimenter by behaving in a way that would make the experiment come out as the participants think it should. Both of these outcomes are disastrous for the experimenter. The experimenter can usually succeed in curbing the participant's desire to be helpful, but the desire to look good is more difficult to curb. Most people do not want to be thought of as weak, abnormal, conformist, unattractive, stupid, or crazy. Thus, if given a chance to figure out what the experimenter is looking for, most people will try to make themselves look good or normal. For example, in an experiment designed specifically

to elucidate this phenomenon,[10] when we told participants that a particular outcome indicated they possessed a good personality trait, they exhibited the behavior necessary to produce that outcome far more often than when we told them it reflected a negative trait. Although this behavior is understandable, it does interfere with meaningful results. For this reason, experimenters find it necessary to deceive participants about the true nature of the experiment.

To illustrate, let's look again at Solomon Asch's[11] classic experiment on conformity. Recall that, in this study, a student was assigned the task of judging the relative size of a few lines. It was a simple task. But a few other students (who were actually accomplices of the experimenter) purposely stated an incorrect judgment. When faced with this situation, a sizable number of the participants yielded to the implicit group pressure and stated an incorrect judgment. This was, of course, a highly deceptive experiment. The participants thought they were participating in an experiment on perception, but, actually, their conformity was being studied. Was this deception necessary? I think so. Let's play it back without the deception: Imagine yourself being a participant in an experiment in which the experimenter said, "I am interested in studying whether or not you will conform in the face of group pressure," and then he told you what was going to happen. My guess is that you wouldn't conform. My guess is that almost *no one* would conform—because conformity is considered a weak and unattractive behavior. What could the experimenter have concluded from this? That people tend to be nonconformists? Such a conclusion would be erroneous and misleading. Such an experiment would be meaningless.

Recall Milgram's experiments on obedience. He found that around 65 percent of the average citizens in his experiment were willing to administer intense shocks to another person in obedience to the experimenter's command. Yet, each year, when I describe the experimental situation to the students in my class and ask them if *they* would obey such a command, only 1 percent indicate that they would. Does this mean my students are nicer people than Milgram's participants? I don't think so. I think it means that people, if given half a chance, will try to look good. Thus, unless Milgram had used deception, he would have come out with results that simply do not reflect the way people behave when they are led to believe they are in real situations. If we were to give people the opportunity to sit

back, relax, and make a guess as to how they would behave in a certain situation, we would get a picture of how people would like to be rather than a picture of how people are.

Ethical Problems

Using deception may be the best (and perhaps the *only*) way to get useful information about the way people behave in most complex and important situations, but it *does* present the experimenter with serious ethical problems. Basically, there are three problems.

1. It is simply unethical to tell lies to people. This takes on even greater significance in the post-Watergate era, when it has been revealed that government agencies have bugged citizens illegally; that presidents tell outright lies to the people who elected them; and that all manner of dirty tricks, fake letters, forged documents, and so on have been used by people directly employed by the president. Can social scientists justify adding to the pollution of deception that currently exists?
2. Such deception frequently leads to an invasion of privacy. When participants do not know what the experimenter is really studying, they are in no position to give their informed consent. For example, in Asch's experiment, it is conceivable that some students might not have agreed to participate had they known in advance that Asch was interested in examining their tendency toward conformity rather than their perceptual judgment.
3. Experimental procedures often entail some unpleasant experiences, such as pain, boredom, anxiety, and the like.

I hasten to add that ethical problems arise even when deception is not used and when experimental procedures are not extreme. Sometimes even the most seemingly benign procedure can profoundly affect a few participants in ways that could not easily have been anticipated—even by the most sensitive and caring experimenters. Consider a series of experiments conducted by Robyn Dawes, Jeanne McTavish, and Harriet Shaklee.[12] Typically, in their investigations of "social dilemmas," participants are faced with the decision to cooperate or to "defect." If everyone cooperates, everyone benefits financially; but if one or more participants choose to

defect, they receive a high payoff, and those who choose to cooperate are at a financial disadvantage. Responses are anonymous and remain so throughout the course of the study. The rules of the game are fully explained to all participants at the beginning of the experiment. And no deception is involved. This scenario seems innocuous enough.

But twenty-four hours after one experimental session, an elderly man telephoned the experimenter. He had been the only defector in his group and had won $190. He wanted to return his winnings and have them divided among the other participants (who had cooperated and won only $1 each). During the conversation, he revealed that he felt miserable about his greedy behavior, that he hadn't slept all night, and so on. After a similar experiment, a woman who cooperated while others defected reported that she felt gullible and had learned that people were not as trustworthy as she had earlier believed.

Despite careful planning by the investigators, the experiments had a powerful impact on participants *that could not have been easily anticipated.* I intentionally chose the experiments by Dawes, McTavish, and Shaklee because they involved no deception and were well within the bounds of ethical codes. My point is simple, but important: No code of ethics can anticipate all problems, especially those created when participants discover something unpleasant about themselves or others in the course of their participation.

Social psychologists who conduct experiments are deeply concerned about ethical issues—precisely because their work is constructed on an ethical dilemma. Let me explain. This dilemma is based on two conflicting values to which most social psychologists subscribe. On the one hand, they believe in the value of free scientific inquiry. On the other hand, they believe in the dignity of humans and their right to privacy. This dilemma is a real one and cannot be dismissed either by piously defending the importance of preserving human dignity or by glibly pledging allegiance to the cause of science. And social psychologists must face this problem squarely, not just once, but each and every time they design and conduct an experiment—for there is no concrete and universal set of rules or guidelines capable of governing *every* experiment.

Obviously, some experimental techniques present more problems than others. In general, experiments that employ deception are

cause for concern because the act of lying is, *in itself,* objectionable—even if the deception is at the service of uncovering the truth. And procedures that cause pain, embarrassment, guilt, or other intense feelings present obvious ethical problems.

More subtle but no less important ethical problems result when participants confront some aspect of themselves that is not pleasant or positive. Recall the experiences of the participants in the relatively mild experiments by Dawes, McTavish, and Shaklee. And many of Solomon Asch's[13] participants learned that they would conform in the face of group pressure; many participants in our own experiment (Aronson and Mettee)[14] learned that they were capable of cheating at a game of cards; most of Milgram's[15] participants learned that they would obey an authority even if such obedience (apparently) involved harming another person.

It could be argued that such self-discovery is of therapeutic or educational benefit to participants; indeed, many participants themselves have made this point. But this does not, in itself, justify these procedures. After all, how could an experimenter know in advance that it would be therapeutic? Morever, it is arrogant of any scientist to decide that he or she has the right or the skill to provide people with a therapeutic experience without their prior permission to do so.

Given these problems, do the ends of social psychological research justify the means? This is a debatable point. Some argue that, no matter what the goals of this science are and no matter what the accomplishments, they are not worth it if people are deceived or put through some discomfort. On the opposite end of the spectrum, others insist that social psychologists are finding things out that may have profound benefits for humankind, and accordingly, almost any price is worth paying for the results.

My own position is somewhere in between. I believe the science of social psychology is important, and I also believe that the health and welfare of experimental participants should be protected at all times. When deciding whether a particular experimental procedure is ethical, I believe a cost-benefit analysis is appropriate. That is, we should consider how much good will derive from doing the experiment and how much harm will be done to the experimental participants. Put another way, the benefits to science and society are compared with the costs to the participants, and this ratio is entered

into the decision calculus. Unfortunately, such a comparison is difficult to make because we can never be absolutely certain of either the benefit or the harm in advance of the experiment.

Consider the obedience experiment. On the face of it, it was a difficult procedure, all right—no doubt about it. But Milgram had no way of knowing exactly *how* difficult it was until he was deeply into the experiment. In my opinion, it was also an extremely important experiment; it taught us a great deal about human behavior. In the balance, I'm glad that Milgram went ahead with it. Not everyone will agree with me. Immediately after its publication, the experiment was lambasted on ethical grounds, both by the popular press and by serious scientists. A few years after having published his results, Stanley Milgram confided in me—sadly, and with a tinge of bitterness—that he believed much of the criticism was fueled by the results he obtained rather than by the actual procedure he employed. That, in and of itself, is an interesting question: Would the criticisms of the ethics of Milgram's procedure have been less vehement if none of the participants had administered shocks beyond a moderate level of intensity? More than a decade later, Leonard Bickman and Matthew Zarantonello[16] discovered that Milgram's ruminations were on target. They did a simple little experiment in which they asked 100 people to read the procedure section of Milgram's experiment. Those people who were informed that a high proportion of Milgram's participants had been fully obedient rated the procedure as more harmful (and, therefore, less ethical) than those who were informed that hardly anyone had been fully obedient. On a more general note, I would suggest that the ethics of any experiment would seem less problematic when the results tell us something pleasant or flattering about human nature than when they tell us something we'd rather not know. That certainly doesn't mean that we should limit our research to the discovery of flattering things! Milgram's obedience experiment is an excellent case in point. I believe that, if a scientist is interested in studying the extent to which a person will harm others in blind obedience to authority, there is no way of doing it without producing some degree of discomfort.

In sum, a social psychologist's decision whether to do a particular experiment depends on an assessment of the potential costs and benefits of that specific experiment. When my students are

contemplating whether to go forward with an experiment, I advise them to use the following five guidelines.

1. Procedures that cause intense pain or intense discomfort should be avoided, if at all possible. Depending on the hypothesis being tested, some discomfort may be unavoidable.
2. Experimenters should provide their participants with the real option of quitting the experiment if their discomfort becomes too intense.
3. Experimenters should be alert to alternative procedures to deception. If some other viable procedure can be found, it should be used.
4. Experimenters should spend considerable time with each participant at the close of the experimental session, carefully explaining the details of the experiment, its true purpose, the reasons for the deception or discomfort, and so on. During this "debriefing" session, they should go out of their way to protect the dignity of participants, to avoid making them feel stupid or gullible about having "fallen for" the deception. They should make certain that participants leave the scene in good spirits—feeling good about themselves and their role in the experiment. This can be accomplished by any earnest experimenter who is willing to put in the time and effort to repay each participant (with information and consideration) for the important role that he or she has played in the scientific enterprise.
5. Finally, experimenters should not undertake an experiment that employs deception or discomfort "just for the hell of it." Before entering the laboratory, experimenters should be certain their experiment is sound and worthwhile—that they are seeking the answer to an interesting question and doing so in a careful, well-organized manner.

Experimenters in social psychology try hard to be as sensitive as possible to the needs of their participants. Although many experiments involve procedures that cause some degree of discomfort, the vast majority of these procedures contain many safeguards for the protection of participants. Again, let us return to the obedience experiment simply because, from the perspective of the participants, it is among the most stressful procedures reported in this book. It is

evident that Milgram worked hard after the experiment to turn the overall experience into a useful and exciting one for his participants. It is also clear that his efforts achieved a high degree of success: Several weeks after the experiment, 84 percent of the participants reported that they were glad to have taken part in the study; 15 percent reported neutral feelings; and only 1 percent stated that they were sorry they had participated. (We should view these findings with caution, however. The discussion of cognitive dissonance in Chapter 5 has taught us that people sometimes justify their behavior by changing their previously held attitudes.) More convincing evidence comes from a follow-up study: One year after the experimental program was completed, a university psychiatrist interviewed a random sample of the participants and found no evidence of injurious effects; rather, the typical response was that their participation was instructive and enriching.[17]

Our Debt to Participants In this chapter, I have discussed the advantages of the experimental method and have shown how complex and challenging it is to design a laboratory experiment in social psychology. In addition, I have shared some of the excitement I feel in overcoming difficulties and discussed ways of ensuring the well-being, as well as the learning, of the participants in our experiments. The knowledge, information, and insights into human social behavior described in the first eight chapters of this book are based on the techniques and procedures discussed in this chapter. They are also based on the cooperation of tens of thousands of individuals who have allowed us to study their behavior in laboratories all over the world. We owe them a lot. Ultimately, our understanding of human beings in all their complexity rests on our ingenuity in developing techniques for studying behavior that are well controlled and influential without violating the essential dignity of those individuals who contribute to our understanding by serving as experimental participants.

What If Our Discoveries Are Misused?

There is one additional ethical consideration: the moral responsibility of the scientist for what he or she discovers. Throughout this book, I have been dealing with some powerful antecedents of persuasion. This was particularly true in Chapter 5, where I discussed techniques

of inducing self-persuasion, and in some of the subsequent chapters, where I discussed applications of these techniques. Self-persuasion is a very powerful force because, in a very real sense, the persuaded never know what hit them. They come to believe that a particular thing is true, not because J. Robert Oppenheimer or T. S. Eliot or Joe "The Shoulder" convinced them it is true, but because they have convinced *themselves*. What's more, they frequently do not know why or how they came to believe it. This renders the phenomenon not only powerful, but frightening, as well. As long as I know why I came to believe X, I am relatively free to change my mind; but if all I know is that X is true—and that's all there is to it—I am far more likely to cling to that belief, even in the face of a barrage of disconfirming evidence.

The mechanisms I have described can be used to get people to floss their teeth, to stop bullying smaller people, to reduce pain, or to love their neighbors. Many people might consider these good outcomes, but they are manipulative just the same. Moreover, the same mechanisms can also be used to get people to buy particular brands of toothpaste and perhaps to vote for particular political candidates. In this era of political spin doctors, propagandists, and hucksters, isn't it immoral to use powerful techniques of social influence?

As the reader of this volume must know by this time, as a real person living in the real world, I have many values—and have made no effort to conceal them; they stick out all over the place. For example, I would like to eliminate bigotry and cruelty. If I had the power, I would employ the most humane and effective methods at my disposal to achieve those ends. I am equally aware that, once these methods are developed, others might use them to achieve ends I might not agree with. This causes me great concern. I am also aware that you may not share my values. Therefore, if you believe these techniques are powerful, you should be concerned.

At the same time, I hasten to point out that the phenomena I have been describing on these pages are not entirely new. After all, it was not a social psychologist who got Mr. Landry hooked on Marlboros, or who invented low-balling; and it was not a social psychologist who induced Lieutenant Calley to attempt to justify the wanton killing of Vietnamese civilians. They did what they did on their own. Social psychologists are attempting to understand these phenomena and scores of others that take place in the world every day—some of

which have been occurring since the first two people on earth began interacting. By understanding these phenomena, the social psychologist may be able to help people understand the processes and consequences involved and possibly refrain from performing a particular behavior when they themselves decide it is dysfunctional.

But the mere fact that we, as working social psychologists, know that the phenomena we deal with are not of our own creation does not free us from moral responsibility. Our research often crystallizes these phenomena into highly structured, easily applicable techniques. There is always the possibility that some individuals may develop these techniques and use them for their own ends. In the hands of a demagogue, these techniques could conceivably turn our society into an Orwellian nightmare. It is not my intention to preach about the responsibilities of social psychologists. What I am most cognizant of are what I believe to be my own responsibilities. Briefly, they are to educate the public about how these techniques might be used and to remain vigilant against their abuse as I continue to do research aimed at furthering our understanding of us social animals—how we think, how we behave, what makes us aggressive, and what makes us loving. Frankly, I can think of no endeavor more interesting or more important.