9

Conducting Experiments

- **Selecting Research Participants**
- **Manipulating the Independent Variable**
  - Setting the Stage
  - Types of Manipulations
    - *Straightforward Manipulations*
    - *Staged Manipulations*
  - Strength of the Manipulation
  - Cost of the Manipulation
- **Measuring the Dependent Variable**
  - Types of Measures
    - *Self-Report Measures*
    - *Behavioral Measures*
    - *Physiological Measures*
  - Sensitivity of the Dependent Variable
  - Multiple Measures
  - Cost of Measures
  - Ethics
- **Additional Controls**
  - Controlling for Participant Expectations
    - *Demand Characteristics*
    - *Placebo Groups*
  - Controlling for Experimenter Expectations
    - *Research on Expectancy Effects*
    - *Solutions to the Expectancy Problem*
- **Debugging the Study**
  - Research Proposals
  - Pilot Studies
  - Manipulation Checks
  - Debriefing
- **Using Computers to Conduct Research**
- **Analyzing and Interpreting Results**
- **Communicating Research to Others**
  - Professional Meetings
  - Journal Articles

Study Terms

Review Questions

Activity Questions

The previous chapters have laid the foundation for planning a research investigation. In this chapter, we will focus on some very practical aspects of conducting research. How do you select the research participants? What should you consider when deciding how to manipulate an independent variable? What should you worry about when you measure a variable? What do you do when the study is completed?

SELECTING RESEARCH PARTICIPANTS

The focus of your study may be children, college students, schizophrenics, rats, pigeons, rabbits, primates, or even cockroaches or flatworms; in all cases, the participants or subjects must somehow be selected. The method used to select participants has implications for generalizing the research results.

Recall from Chapter 7 that most research projects involve sampling research participants from a population of interest. The population is composed of all of the individuals of interest to the researcher. Samples may be drawn from the population using probability sampling or nonprobability sampling techniques. When it is important to accurately describe the population, you must use probability sampling. This is why probability sampling is so crucial when conducting scientific polls. Much research, however, is more interested in testing hypotheses about behavior. Here, the focus of the study is the relationships between the variables being studied and testing predictions derived from theories of behavior. In such cases, the participants may be found in the easiest way possible using nonprobability haphazard or "convenience" sampling methods. You may ask students in introductory psychology classes to participate, knock on doors in your dorm to find people to be tested, or choose a class in which to test children simply because you know the teacher. Nothing is wrong with such methods as long as you recognize that they affect the ability to generalize your results to some larger population. The issue of generalizing results is discussed in Chapter 14; despite the problems of generalizing results based on convenient haphazard samples, ample evidence supports the view that we *can* generalize findings to other populations and situations.

You will also need to determine your sample size. How many participants will you need in your study? In general, increasing your sample size increases the likelihood that your results will be statistically significant, because larger samples provide more accurate estimates of population values (see Table 7.1). Most researchers pay attention to the sample sizes in the research area being studied and select a sample size that is typical for studies in the area. A more formal approach to selecting a sample size is discussed in Chapter 13.

MANIPULATING THE INDEPENDENT VARIABLE

To manipulate an independent variable, you have to construct an operational variable into a set of operations—specific instructions, events, and stimuli to be presented to the research participants. In addition, the independent and dependent variables must be introduced within the context of the total experimental setting. This has been called "setting the stage" (Aronson, Brewer, & Carlsmith, 1985).

Setting the Stage

In setting the stage, you usually have to do two things: provide the participants with the informed consent information needed for your study and explain to participants why the experiment is being conducted. Sometimes, the rationale given is completely truthful, although only rarely will you want to tell participants the actual hypothesis. For example, you might say that you are conducting an experiment on memory when, in fact, you are studying a specific aspect of memory (your independent variable). If participants know what you are studying, they may try to confirm the hypothesis, or they may try to look good by behaving in the most socially acceptable way. If you find that deception is necessary, you have a special obligation to address the deception when you debrief the participants at the conclusion of the experiment.

There are no clear-cut rules for setting the stage, except that the experimental setting must seem plausible to the participants, nor are there any clear-cut rules for translating conceptual variables into specific operations. Exactly how the variable is manipulated depends on the variable and the cost, practicality, and ethics of the procedures being considered.

Types of Manipulations

Straightforward manipulations Researchers are usually able to manipulate a variable with relative simplicity by presenting written, verbal, or visual material to the participants. Such *straightforward* manipulations manipulate variables with instructions and stimulus presentations. Let's look at a few examples.

Labranche, Helweg-Larsen, Byrd, and Choquette (1997) studied the impact of health promotion brochures by asking women to read a brochure on breast self-examinations. In one condition, the brochure included only text; in the other condition, pictures depicting breast self-examination were added to the brochure. Participants' responses to the two brochures depended on their level of comfort with sexual materials. One question asked about whether the woman believed she could properly perform a breast self-examination. Women who were uncomfortable with sexual materials were less sure about their ability when they read the brochure with pictures than when they read the text-only brochure. The type of brochure did not affect the women who were comfortable with sexual materials.

Studies on jury decisions often ask participants to read a description of a jury trial in which a crucial piece of information is varied. Bornstein (1998) studied the effect of the severity of injury on product liability judgments. Participants read about a case in which a woman taking birth-control pills had been diagnosed with

cancer. In a low-severity condition, the cancer was detected early, one ovary was removed, the woman could still have children, and future prognosis was good. In the high-severity condition, the cancer was detected late, both ovaries were removed so pregnancy would not be possible, and the future prognosis was poor. The evidence on whether the pills could be responsible for the cancer was the same in both conditions; thus, product reliability judgments should be the same in both conditions. Nevertheless, the severity information affected liability judgments: the pill manufacturer was found liable by 40% of the participants in the high-severity condition versus 21% in the low-severity condition.

Most memory research relies on straightforward manipulations. For example, Coltheart and Langdon (1998) displayed lists of words to participants and later measured recall. The word lists differed on phonological similarity: Some lists had words that sounded similar, such as *cat, map,* and *pat,* and other lists had dissimilar words such as *mop, pen,* and *cow.* They found that lists with dissimilar words are recalled more accurately. In a more complex memory study, Reeve and Aggleton (1998) presented a script of a future episode of a British soap opera called "The Archers" to both fans ("experts") and people unfamiliar with the show. In one condition, the script was typical of an actual episode of the program—the Archers visit a livestock market. In the other condition, the script was atypical—the Archers visit a boat show. The characters and basic structure of the show were identical in the two conditions. After reading the script, the participants were given a measure of retention of the details of the episode. They found that being an expert aided retention only when the story was a typical one. In the atypical condition, both fans and non-fans had equal retention. Reeve and Aggleton concluded that the benefits of being an expert are very limited.

As a final example of a straightforward manipulation, consider a study by Petty, Cacioppo, and Goldman (1981) on the effect of communicator credibility and personal involvement on attitude change. The participants were college seniors who read about the reasons that a comprehensive examination should be required for graduation from their university. To manipulate credibility, the arguments were said to be written by either a professor of education at Princeton University or a junior at a local school. The researchers also manipulated personal involvement by telling the students that the examination was being considered for implementation either that year (thus affecting the individuals participating in the study) or 10 years later. Participants in the low-involvement condition changed their attitudes more if the communicator was high in credibility, but the credibility of the communicator did not make a difference when the participants were highly involved.

You will find that most manipulations of independent variables in many areas of research are straightforward. Researchers vary the difficulty of material to be learned, motivation levels, the way questions are asked, characteristics of people to be judged, and a variety of other factors in a straightforward manner.

Staged manipulations Other manipulations are less straightforward. Sometimes, it is necessary to stage events that occur during the independent variable successfully. When this occurs, the manipulation is called a *staged* or *event manipulation.*

Staged manipulations are most frequently used for two reasons. First, the researcher may be trying to create some psychological state in the participants, such as frustration or a temporary lowering of self-esteem; second, a staged manipulation may be necessary to simulate some situation that occurs in the real world. For example, Fazio, Cooper, Dayson, and Johnson (1981) studied cognitive performance under conditions of multiple task demands. Participants in one condition spent 10 minutes proofreading a manuscript; participants in the other condition performed the same proofreading task but were interrupted by the experimenter from time to time and asked to go to another room to perform other tasks. These conditions simulate common real-world work environments.

Staged manipulations frequently employ a **confederate** (sometimes termed an "accomplice"). Usually, the confederate appears to be another participant in an experiment but is actually part of the manipulation. We discussed the use of confederates in Chapter 3. For example, in a study on aggression, the confederate and the participant both report to the experiment and are told to wait in a room for the experiment to begin. During the waiting period, the confederate insults the participant in an "anger" condition but does not insult the participant in a "no-anger" condition. The experimenter then enters and informs the two individuals that learning is being studied and that one of them will be a teacher and the other will be a learner. The assignments to the roles of teacher and learner appear to be random but are actually rigged by the experimenter: The confederate is always the learner and the real participant is always the teacher. In the learning task, the participant is permitted to shock the confederate whenever an incorrect answer is given. The amount of shock that is chosen is the measure of aggression; the researcher compares the amount of shock given in the anger and no-anger conditions. Confederates are also used in field experiments; for example, an accomplice may appear to be merely another person in the setting, such as a shopper at a mall who asks you for change (Baron, 1997).

Staged manipulations demand a great deal of ingenuity and even some acting ability. They are used to involve the participants in an ongoing social situation, which the individuals perceive not as an experiment but as a real experience. Researchers assume that the result will be natural behavior that truly reflects the feelings and intentions of the participants. However, such procedures allow for a great deal of subtle interpersonal communication that is hard to put into words; this may make it difficult for other researchers to replicate the experiment. Also, a complex manipulation is difficult to interpret. If many things happened during the experiment, what *one* thing was responsible for the results? In general, it is easier to interpret results when the manipulation is relatively straightforward. However, the nature of the variable you are studying sometimes demands complicated procedures.

Strength of the Manipulation

The simplest experimental design has two levels of the independent variable. In planning the experiment, the researcher has to choose these levels. A general principle to follow is to make the manipulation as strong as possible. A strong manipulation maximizes the differences between the two groups and increases the chances that the independent variable will have a statistically significant effect on the dependent variable.

To illustrate, suppose you think that there is a positive linear relationship between attitude similarity and liking ("birds of a feather flock together"). In conducting the experiment, you could arrange for participants to encounter another person, a confederate. In one group, the confederate and the participant would share similar attitudes; in the other group, the confederate and the participant would be dissimilar. Similarity, then, is the independent variable, and liking is the dependent variable. Now you have to decide on the amount of similarity. Figure 9.1 shows the hypothesized relationship between attitude similarity and liking at ten different levels of similarity. Level 1 represents the least amount of similarity with no common attitudes, and level 10 the greatest (all attitudes are similar). To achieve the strongest manipulation, the participants in one group would encounter a confederate of level 1 similarity and those in the other group would encounter a confederate of level 10 similarity. This would result in the greatest difference in the liking means—a 9-point difference. A weaker manipulation—using levels 4 and 7, for example—would result in a smaller mean difference.

A strong manipulation is particularly important in the early stages of research, when the researcher is most interested in demonstrating that a relationship does, in fact, exist. If the early experiments reveal a relationship between the

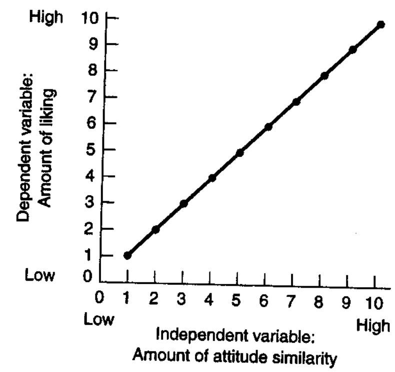

FIGURE 9.1
Relationship between attitude similarity and liking

variables, subsequent research can systematically manipulate the other levels of the independent variable to provide a more detailed picture of the relationship.

The principle of using the strongest manipulation possible should be tempered by at least two considerations. First, the strongest possible manipulation may involve a situation that rarely, if ever, occurs in the real world. For example, an extremely strong crowding manipulation might involve placing so many people in a room that no one could move – a manipulation that might significantly affect a variety of behaviors. However, we wouldn't know if the results were similar to those occurring in more common, less crowded situations such as many classrooms or offices.

A second consideration is ethics: A manipulation should be as strong as possible within the bounds of ethics. A strong manipulation of fear or anxiety, for example, might not be possible because of the potential physical and psychological harm to participants.

Cost of the Manipulation

Cost is another factor in the decision about how to manipulate the independent variable. Researchers who have limited monetary resources may not be able to afford expensive equipment, salaries for confederates, or payments to participants in long-term experiments. Also, a manipulation in which participants must be run individually requires more of the researcher's time than a manipulation that allows running many individuals in a single setting. In this respect, a manipulation that uses straightforward presentation of written or verbal material is less costly than a complex, staged, experimental manipulation. Some government and private agencies offer grants for research; because much research is costly, continued public support of these agencies is very important.

MEASURING THE DEPENDENT VARIABLE

In previous chapters, we have discussed various aspects of measuring variables including reliability, validity, and reactivity of measures, observational methods, and the development of self-report measures for questionnaires and interviews. In this chapter, we will focus on measurement considerations that are particularly relevant to experimental research.

Types of Measures

The dependent variable in most experiments is one of three general types: self-report, behavioral, or physiological.

Self-report measures Self-reports can be used to measure attitudes, liking for someone, judgments about someone's personality characteristics, intended behaviors, emotional states, attributions about why someone performed well or poorly on a task, confidence in one's judgments, and many other aspects of

human thought and behavior. Rating scales with descriptive anchors (end points) are most commonly used. For example, the Labranche et al. (1997) study described earlier asked women to respond on a 7-point scale after they read the brochure:

I feel I could properly give myself a breast self-examination.

Strongly disagree ____ ____ ____ ____ ____ ____ ____ Strongly agree

Behavioral measures Behavioral measures are direct observations of behaviors. As with self-reports, measurements of an almost endless number of behaviors are possible. Sometimes, the researcher may record whether or not a given behavior occurs – for example, whether or not an individual responds to a request for help, makes an error on a test, or chooses to engage in one activity rather than another. Often, the researcher must decide whether to record the number of times a behavior occurs in a given time period – the *rate* of a behavior; how quickly a response occurs after a stimulus – a *reaction time;* or how long a behavior lasts – a measure of *duration.* The decision of which aspect of behavior to measure depends on which is most theoretically relevant for the study of a particular problem or which measure logically follows from the independent variable manipulation.

Sometimes, the nature of the variable being studied requires either a self-report or behavioral measure. A measure of helping behavior is almost by definition a behavioral measure, whereas a measure of perception of the personality characteristics of someone will employ a self-report measure. For many variables, however, both self-reports and behavioral measures could be appropriate. Thus, liking or attraction could be measured on a rating scale or with a behavioral measure of the distance two people place between themselves or the amount of time they spend looking into each other's eyes. When both options are possible, a series of studies may be conducted to study the effects of an independent variable on both types of measures.

Physiological measures Physiological measures are recordings of the physiological responses of the body. Many such responses are available; examples include the **galvanic skin response** (GSR), **electromyogram** (EMG), and **electroencephalogram** (EEG). The GSR is a measure of general emotional arousal and anxiety; it measures the electrical conductance of the skin, which changes when sweating occurs. The EMG measures muscle tension and is frequently used as a measure of tension or stress. The EEG is a measure of electrical activity of brain cells. It can be used to record general brain arousal as a response to different situations, activity in different parts of the brain as learning occurs, or brain activity during different stages of sleep.

The GSR, EMG, and EEG have long been used as physiological indicators of important psychological variables. Many oth available, including temperature, heart rate, and information that can be gathered from blood or urine analysis. Often, such measures offer valuable alternatives to self-report and behavioral measures (also see Cacioppo & Tassinary, 1990).

Sensitivity of the Dependent Variable

The dependent variable should be sensitive enough to detect differences between groups. A measure of liking that asks, "Do you like this person?" with a simple "yes" or "no" response alternative is less sensitive than one that asks, "How much do you like this person?" on a 5- or 7-point scale. With the first measure, people may tend to be nice and say yes even if they have some negative feelings about the person. The second measure allows for a gradation of liking; such a scale would make it easier to detect differences in the amount of liking.

The issue of **sensitivity** is particularly important when measuring human performance. Memory can be measured using recall, recognition, or reaction time; cognitive task performance might be measured by examining speed or number of errors during a proofreading task; physical performance can be measured through various motor tasks. Such tasks vary in their difficulty. Sometimes a task is so easy that everyone does well regardless of the conditions that are manipulated by the independent variable. This results in what is called a **ceiling effect**—the independent variable appears to have no effect on the dependent measure only because participants quickly reach the maximum performance level. The opposite problem occurs when a task is so difficult that hardly anyone can perform well; this is called a **floor effect.**

The need to consider sensitivity of measures is nicely illustrated in the Freedman et al. (1971) study of crowding mentioned in Chapter 4. The study examined the effect of crowding on various measures of cognitive task performance and found that crowding did not impair performance. You could conclude that crowding has no effect on performance; however, it is also possible that the measures were either too easy or too difficult to detect an effect of crowding. In fact, subsequent research showed that the tasks may have been too easy; when participants were asked to perform more complex tasks, crowding did result in lower performance (Paulus, Annis, Seta, Schkade, & Matthews, 1976).

Multiple Measures

It is often desirable to measure more than one dependent variable. One reason to use multiple measures stems from the fact that a variable can be measured in a variety of concrete ways (recall the discussion of operational definitions in Chapter 4). In a study on health-related behaviors, for example, researchers measured the number of work days missed because of ill health, the number of doctor visits, and the use of aspirin and tranquilizers (Matteson & Ivancevich, 1983). Physiological measures might have been taken as well. If the independent variable has the same effect on several measures of the same dependent variable, our confidence in the results is increased. It is also useful to know whether the same independent variable affects some measures but not others. For example, an

independent variable designed to affect liking might have an effect on some measures of liking (e.g., desirability as a person to work with) but not others (e.g., desirability as a dating partner). Researchers also may be interested in studying the effects of an independent variable on several different behaviors. For example, an experiment on the effects of a new classroom management technique might examine academic performance, interaction rates among classmates, and teacher satisfaction.

Making multiple measurements in a single experiment is valuable when it is feasible to do so. However, it may be necessary to conduct a series of experiments to explore the effects of an independent variable on various behaviors.

Cost of Measures

Another consideration is cost—some measures may be more costly than others. Paper-and-pencil self-report measures are generally inexpensive; measures that require trained observers or elaborate equipment can become quite costly. A researcher studying nonverbal behavior, for example, might have to use a video camera to record each participant's behaviors in a situation. Two or more observers would then have to view the tapes to code behaviors such as eye contact, smiling, or self-touching (two observers are needed to ensure that the observations are reliable). Thus, there would be expenses for both equipment and personnel. Physiological recording devices are also expensive. Researchers need resources from the university or outside agencies to carry out such research.

Ethics

Ethical concerns are always important. Researchers must be extremely careful about potential invasion of privacy and must always ensure that responses are completely confidential.

ADDITIONAL CONTROLS

The basic experimental design has two groups: in the simplest case, an experimental group that receives the manipulation and a control group that does not. Use of a control group makes it possible to eliminate a variety of alternative explanations based on history, maturation, statistical regression, and so on. Sometimes additional control procedures may be necessary to address other types of alternative explanations. Two general control issues concern expectancies on the part of both the participants in the experiment and the experimenters.

Controlling for Participant Expectations

Demand characteristics We noted previously that experimenters do not wish to inform participants about the specific hypotheses being studied or the exact purpose of the research. The reason for this lies in the probl...