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 non-probability 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 experimental...
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 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 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, Reese 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. Reese 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 article was 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 a study.
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 partic-
ipant 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 con-
 federate 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 an
other 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 ac-
ing 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 experi-
ment. Also, a complex manipulation is difficult to interpret. If many things hap-
pened 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 ef-
fect 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 an-
other person, a confederate. In one group, the confederate and the participant
would share similar attitudes; in the other group, the confederate and the par-

ticipant would be dissimilar. Similarity, then, is the independent variable, and
liking is the dependent variable. Now you have to decide on the amount of sim-
ilarity. Figure 9.1 shows the hypothesized relationship between attitude simi-
arity 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 dif-
ference. 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 re-
search, when the researcher is most interested in demonstrating that a relation-
ship does, in fact, exist. If the early experiments reveal a relationship between the

![Figure 9.1](image-url)
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 behavior measure of the distance two people place between themselves or the amount of time they spend looking into each others' 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 neurological variables. Many other physiological measures...
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 (Mattson & 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 expectations 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 problem of participant expectations. When participants come to suspect the nature of the research, their behavior may be influenced. For example, if participants are informed of the aims of the experiment, they may try to be consistent with the researchers' expectations. If participants expect to be exposed to some stressful manipulation, for example, they may try to be on the alert for signs of stress in the experiment. Consequently, if this stress occurs, they may interpret it as evidence of the manipulation, thereby increasing the likelihood of a demand characteristic. Therefore, it is important to keep participants in the dark about the specific hypotheses being studied so that they are not consciously manipulated.
characteristics (Orne, 1962). A demand characteristic is any feature of an experiment that might inform participants of the purpose of the study. The concern is that when participants form expectations about the hypothesis of the study, they will then do whatever is necessary to confirm the hypothesis; this, of course, assumes they are motivated to be cooperative. Orne conducted research to demonstrate that people are in fact cooperative. For example, he asked participants to add numbers on a sheet of paper; when they had finished, they picked up a card from a large stack for further instructions. Each instruction card said to tear the sheet into 32 pieces and to go to the next page of numbers. The participants continued this ridiculous task for several hours with no protest or questioning! Although you can probably think of situations in which the individuals would try to be uncooperative, Orne’s conception of the cooperative participant seems to be generally correct.

One way to control for demand characteristics is to use deception— to make participants think that the experiment is studying one thing when actually it is studying something else. The experimenter may devise elaborate cover stories to explain the purpose of the study and to disguise what is really being studied. The researcher may also attempt to disguise the dependent measure by using an unobtrusive measure or by placing the measure among a set of unrelated filler items on a questionnaire. Another approach is simply to assess whether demand characteristics are a problem by asking participants about their perceptions of the purpose of the research. It may be that participants do not have an accurate view of the purpose of the study—or if some individuals do guess the hypotheses of the study, their data may be analyzed separately.

Demand characteristics may be eliminated when people are not aware that an experiment is taking place or that their behavior is being observed. Thus, experiments conducted in field settings and observational research in which the observer is concealed or unobtrusive measures are used minimize the problem of demand characteristics.

Placebo groups A special kind of participant expectation arises in research on the effects of drugs. Consider an experiment that is investigating whether a drug reduces depression in psychiatric patients. One group of people diagnosed as depressive receives the drug and the other group does not. Now suppose that the drug group shows an improvement. We do not know whether the improvement was caused by the properties of the drug or by the expectations about the effect of the drug—what is called a placebo effect. In other words, just administering a pill or an injection may be sufficient to cause an observed improvement in behavior. To control for this possibility, a placebo group can be added. Participants in the placebo group receive a pill or injection containing an inert, harmless substance; they do not receive the drug given to members of the experimental group. If the improvement results from the active properties of the drug, the participants in the experimental group should show greater improvement than those in the placebo group. If the placebo group improves as much as the experimental group, the improvement is a placebo effect.

Sometimes, participants’ expectations are the primary focus of an investigation. For example, Marlatt and Rohsenow (1980) conducted research to determine which behavioral effects of alcohol are due to alcohol itself as opposed to the psychological impact of believing one is drinking alcohol. The experimental design to examine these effects had four groups: (1) expect no alcohol-receive no alcohol, (2) expect no alcohol-receive alcohol, (3) expect alcohol-receive no alcohol, and (4) expect alcohol-receive alcohol. Marlatt and Rohsenow’s research suggests that the belief that one has consumed alcohol is a more important determinant of behavior than the alcohol itself. That is, people who believed they consumed alcohol (Groups 3 and 4) behaved very similarly, although those in Group 3 were not actually given any alcohol.

In some areas of research, the use of control groups has ethical implications. Suppose you are studying a treatment that does have a positive effect on people—for example, by reducing migraine headaches or alleviating symptoms of depression. It is important to use careful experimental procedures to make sure that the treatment does have an impact and that alternative explanations for the effect, including a placebo effect, are eliminated. However, it is also important to help those people who are in the control conditions. Participants in the control conditions may be given the treatment as soon as they have completed their part in the study.

Controlling for Experimenter Expectations

Experimenter are usually aware of the purpose of the study and thus may develop expectations about how participants should respond. These expectations can in turn bias the results. This general problem is called experimenter bias or expectancy effects (Rosenthal, 1966, 1967, 1969).

Expectancy effects may occur whenever the experimenter knows which condition the participants are in. There are two potential sources of experimenter bias. First, the experimenter might unintentionally treat participants differently in the various conditions of the study. For example, certain words might be emphasized when reading instructions to one group but not the other, or the experimenter might smile more when interacting with people in one of the conditions. The second source of bias can occur when experimenters record the behaviors of the participants; there may be subtle differences in the way the experimenter interprets and records the behaviors.

Research on expectancy effects Expectancy effects have been studied in a variety of ways. Perhaps the earliest demonstration of the problem is the case of Clever Hans, a horse whose alleged brilliance was revealed by Pfungst (1911) to be an illusion. Robert Rosenthal describes Clever Hans:

Hans, it will be remembered, was the clever horse who could solve problems of mathematics and musical harmony with equal skill and grace, simply by tapping out the answers with his hoof. A committee of eminent experts...
tested that Hans, whose owner made no profit from his horse's talents, was receiving no cues from his questioner. Of course, Pfungst later showed that this was not so, that tiny head and eye movements were Hans' signals to begin and to end his tapping. When Hans was asked a question, the questioner looked at Hans' hoof, quite naturally so, for that was the way for him to determine whether Hans' answer was correct. Then, it was discovered that when Hans approached the correct number of taps, the questioner would inadvertently move his head or eyes upward—just enough that Hans could discriminate the cue, but not enough that even trained animal observers or psychologists could see it. 1

If a clever horse can respond to subtle cues, it is reasonable to suppose that clever humans can, too. In fact, research has shown that experimenter expectancies can be communicated to humans by both verbal and nonverbal means (Duncan, Rosenberg, & Finklestein, 1969; Jones & Cooper, 1971).

An example of more systematic research on expectancy effects is a study by Rosenthal (1966). In this experiment, graduate students trained rats that were described as coming from either "maze bright" or "maze dull" genetic strains. The animals actually came from the same strain and had been randomly assigned to the bright and dull categories; however, the "bright" rats did perform better than the "dull" rats. Subtle differences in the ways the students treated the rats or recorded their behavior must have caused this result. A generalization of this particular finding is called "teacher expectancy." Research has shown that telling a teacher that a pupil will bloom intellectually over the next year results in an increase in the pupil's IQ score (Rosenthal & Jacobson, 1968). In short, teachers' expectations can influence students' performance.

The problem of expectations influencing ratings of behavior is nicely illustrated in an experiment by Langer and Abelson (1974). Clinical psychologists were shown a videotape of an interview in which the person interviewed was described as either an applicant for a job or a patient; in reality, all saw the same tape. The psychologists later rated the person as more "disturbed" when they thought the person was a patient than when the person was described as a job applicant.

Solutions to the expectancy problem Clearly, experimenter expectations can influence the outcomes of research investigations. How can this problem be solved? Fortunately, there are a number of ways of minimizing expectancy effects. First, experimenters should be well trained and should practice behaving consistently with all participants. The benefit of training was illustrated in the Langer and Abelson study with clinical psychologists. The bias of rating the "patient" as disturbed was much less among behavior-oriented therapists than among traditional ones. Presumably, the training of the behavior-oriented therapists led them to focus more on the actual behavior of the person, so they were less influenced by expectations stemming from the label of "patient."

Another solution is to run all conditions simultaneously so that the experimenter's behavior is the same for all participants. This solution is feasible only under certain circumstances, however, such as when the study can be carried out with the use of printed materials or the experimenter's instructions to participants are the same for everyone.

Expectancy effects are also minimized when the procedures are automated. As noted previously, it may be possible to manipulate independent variables and record responses using computers; with automated procedures, the experimenter's expectations are unlikely to influence the results.

A final solution is to use experimenters who are unaware of the hypothesis being investigated. In these cases, the person conducting the study or making observations is blind regarding what is being studied or which condition the participant is in. This procedure originated in drug research using placebo groups. In a single-blind experiment, the participant is unaware of whether a placebo or the actual drug is being administered; in a double-blind experiment, neither the participant nor the experimenter knows whether the placebo or actual treatment is being given. To use a procedure in which the experimenter or observer is unaware of either the hypothesis or the group the participant is in, you must hire other people to conduct the experiment and make observations.

Because researchers are aware of the problem of expectancy effects, solutions such as the ones just described are usually incorporated into the procedures of the study. If a study does have a potential problem of expectancy effects, researchers are bound to notice and will attempt to replicate the experiment with procedures that control for them. The procedures used in scientific research must be precisely defined so they can be replicated by others. This allows other researchers to build on previous research. It is also a self-correcting mechanism that ensures that methodological flaws will be discovered. The importance of replication will be discussed further in Chapter 14.

DEBUGGING THE STUDY

So far, we have discussed several of the factors that a researcher considers when planning a study. Actually conducting the study and analyzing the results is a time-consuming process. Before beginning the research, the investigator wants to be as sure as possible that everything will be done right. How can this be accomplished? There are a number of ways of eliminating the bugs from a study before it starts.

Research Proposals

After putting considerable thought into planning the study, the researcher writes a research proposal. The proposal will have a beginning and an end, and it will

background for the study. The intent is to clearly explain why the research is being done—what questions the research is designed to answer. The details of the procedures that will be used to test the idea are then given. The plans for analysis of the data are also provided. A research proposal is very similar to the introduction and method sections of a journal article.

Such proposals must be included in applications for research grants; ethics review committees require some type of proposal as well (see Chapter 3 for more information on Institutional Review Boards). Preparing a proposal is a good idea in planning any research project. Simply putting your thoughts on paper helps to organize and systematize ideas. In addition, you can show the proposal to friends, colleagues, professors, and other interested parties who can provide useful feedback about the adequacy of your procedures. They may see problems that you didn’t recognize, or they may offer ways of improving the study.

Pilot Studies

When the researcher has finally decided on all the specific aspects of the procedure, it is possible to conduct a pilot study in which the researcher does a “trial run” with a small number of participants. The pilot study will reveal whether participants understand the instructions, whether the total experimental setting seems plausible, whether any confusing questions are being asked, and so on.

Sometimes participants in the pilot study are questioned in detail about the experience following the experiment. Another method is to use the “think aloud” protocol (described in Chapter 7) in which the participants in the pilot study are instructed to verbalize their thoughts about everything that is happening during the study. Such procedures provide the researcher with an opportunity to make any necessary changes in the procedure before doing the entire study. Also, a pilot study allows the experimenters who are collecting the data to become comfortable with their roles and to standardize their procedures.

Manipulation Checks

A manipulation check is an attempt to directly measure whether the independent variable manipulation has the intended effect on the participants. Manipulation checks provide evidence for the construct validity of the manipulation (construct validity was discussed in Chapter 4). If you are manipulating anxiety, for example, a manipulation check will tell you whether participants in the high-anxiety group really were more anxious than those in the low-anxiety condition. The manipulation check might involve a self-report of anxiety, a behavioral measure (such as number of arm and hand movements), or a physiological measure. All manipulation checks, then, ask whether the independent variable manipulation was in fact a successful operationalization of the conceptual variable being studied. Consider, for example, a manipulation of physical attractiveness as an independent variable. In an experiment, participants respond to someone who is supposed to be perceived as attractive or unattractive. The manipulation check in this case would determine whether participants do rate the highly attractive person as more physically attractive.

Manipulation checks are particularly useful in the pilot study to decide whether the independent variable manipulation is in fact having the intended effect. They can also be used in the actual experiment to demonstrate the effectiveness of the manipulation. However, a manipulation check might not be included in the actual experiment if it would distract participants or inform them about the purpose of the experiment.

A manipulation check has two advantages. First, if the check shows that your manipulation was not effective, you have saved the expense of running the actual experiment. You can then turn your attention to changing the manipulation to make it more effective. For instance, if the manipulation check shows that neither the low- nor the high-anxiety group was very anxious, you could change your procedures to increase the anxiety in the high-anxiety condition.

Second, a manipulation check is advantageous if you get nonsignificant results—that is, if the results indicate that no relationship exists between the independent and dependent variables. A manipulation check can identify whether the nonsignificant results are due to a problem in manipulating the independent variable. If your manipulation is not successful, it is only reasonable that you will obtain nonsignificant results. If both groups are equally anxious after you manipulate anxiety, anxiety can’t have any effect on the dependent measure. What if the check shows that the manipulation was successful, but you still get nonsignificant results? Then you know at least that the results were not due to a problem with the manipulation; the reason for not finding a relationship lies elsewhere. Perhaps you had a poor dependent measure, or perhaps there really is no relationship between the variables.

DEBRIEFING

After all the data are collected, a debriefing session is usually held. This is an opportunity for the researcher to interact with the participants to discuss the ethical and educational implications of the study. Debriefing was discussed in detail in Chapter 3 in the context of ethical considerations.

The debriefing session can also provide an opportunity to learn more about what participants were thinking during the experiment. Researchers can ask participants what they believed to be the purpose of the experiment, how they interpreted the independent variable manipulation, and what they were thinking when they responded to the dependent measures. Such information can prove useful in interpreting the results and planning future studies.

Finally, researchers may ask the participants to refrain from discussing the study with others. Such requests are typically made when more people will be participating and they may talk with one another in classes or residence halls. People who have already participated are aware of the general purposes and
procedures; it is important that these individuals not provide expectancies about the study to potential future participants.

USING COMPUTERS TO CONDUCT RESEARCH

It is becoming easier and more common for researchers to use computers as a tool for manipulating independent variables and measuring behaviors. An individual sitting at a computer screen can be presented with written material or graphical displays that replace traditional methods such as printed materials and tachistoscopes (devices that vary the length of presentation of a stimulus), and slides. Researchers can ask questions on a computer monitor instead of using the traditional paper-and-pencil method. Computers can also be used to record response times and control physiological recording devices and other equipment.

More and more researchers are conducting surveys and experiments over the Internet. People connected to the Internet from anywhere on the globe are potential participants in a research investigation. The fact that this is now possible raises many intriguing issues such as the integrity of the data and the nature of the sample. Recall from Chapter 7 that current research findings indicate that data obtained via the Internet are comparable to data obtained with traditional methods.

ANALYZING AND INTERPRETING RESULTS

After the data have been collected, the next step is to analyze them. Statistical analyses of the data are carried out to allow the researcher to examine and interpret the pattern of results obtained in the study. The statistical analysis helps the researcher decide whether there really is a relationship between the independent and dependent variables; the logic underlying the use of statistical tests is discussed in Chapter 13. It is not the purpose of this book to teach statistical methods; however, the calculations involved in several statistical tests are provided in Appendix B.

COMMUNICATING RESEARCH TO OTHERS

The final step is to write a report that details why you conducted the research, how you obtained the participants, what procedures you used, and what you found. A description of how to write such reports is included in Appendix A. After you have written the report, what do you do with it? How do you communicate the findings to others? Research findings are most often submitted as journal articles or as papers to be read at scientific meetings. In either case, the submitted paper is evaluated by two or more knowledgeable reviewers who decide whether the paper is acceptable for publication or presentation at the meeting.

Professional Meetings

Meetings sponsored by professional associations are important opportunities for researchers to present their findings to other researchers and the public. National and regional professional associations such as the American Psychological Association (APA) and the American Psychological Society (APS) hold annual meetings at which psychologists and psychology students present their own research and learn about the latest research being done by their colleagues. Sometimes, verbal presentations are delivered to an audience. However, poster sessions are more common; here, researchers display posters that summarize the research and are available to discuss the research with others.

Journal Articles

As we noted in Chapter 2, there are many journals in which research papers are published. Nevertheless, the number of journals is small compared to the number of reports written; thus, it is not easy to publish research. When a researcher submits a paper to a journal, two or more reviewers read the paper and recommend acceptance (often with the stipulation that revisions be made) or rejection. As many as 75–90% of papers submitted to the more prestigious journals are rejected. Many rejected papers are submitted to other journals and eventually accepted for publication, but much research is never published. This is not necessarily bad; it simply means that selection processes separate high-quality research from that of lesser quality.

Many of the decisions that must be made when planning an experiment were described in this chapter. The discussion focused on experiments that use the simplest experimental design with a single independent variable. In the next chapter, more complex experimental designs are described.

Study Terms

<table>
<thead>
<tr>
<th>Behavioral measure</th>
<th>Manipulation strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ceiling effect</td>
<td>Physiological measure</td>
</tr>
<tr>
<td>Debriefing</td>
<td>Pilot study</td>
</tr>
<tr>
<td>Demand characteristics</td>
<td>Placebo group</td>
</tr>
<tr>
<td>Double-blind procedure</td>
<td>Self-report measure</td>
</tr>
<tr>
<td>Expectancy effects (experimenter bias)</td>
<td>Single-blind procedure</td>
</tr>
<tr>
<td>Floor effect</td>
<td>Staged manipulation</td>
</tr>
<tr>
<td>Manipulation check</td>
<td>Straightforward manipulation</td>
</tr>
</tbody>
</table>
Review Questions

1. Distinguish between staged and straightforward manipulations of an independent variable.
2. Distinguish between the general types of dependent variables.
3. What is meant by the sensitivity of a measure? What are ceiling and floor effects?
4. Discuss the ways that computers can be used in conducting an experiment.
5. What are demand characteristics? Describe ways to minimize demand characteristics.
6. What is the reason for a placebo group?
7. What are experimenter expectancy effects? What are some solutions to the experimenter bias problem?
8. What methods can be used to debug an experiment?
9. What is a pilot study?
10. What is a manipulation check? How does it help the researcher interpret the results of an experiment?
11. What does a researcher do with the findings after completing a research project?

Activity Questions

1. Dr. Romano studied the relationship between age and reading comprehension, specifically predicting that older people will show lower comprehension than younger ones. Romano was particularly interested in comprehension of material that is available in the general press. Groups of participants who were 20, 30, 40, and 50 years of age read a chapter from a book by physicist Stephen W. Hawking (1988) entitled *A Brief History of Time: From the Big Bang to Black Holes* (the book was on the best-seller list at the time). After reading the chapter, participants were given a comprehension measure. The results showed that there was no relationship between age and comprehension scores; all age groups had equally low comprehension scores. Why do you think no relationship was found? Identify at least two possible reasons.

2. Recall the experiment on facilitated communication by children with autism that was described in Chapter 2 (Montee, Miltenberger, & Wittrock, 1995). Interpret the findings of that study in terms of experimenter expectancy effects.

3. Your lab group has been assigned the task of designing an experiment to investigate the effect of time spent studying on a recall task. Thus far, your group has come up with the following plan: "Participants will be randomly assigned to two groups. Individuals in one group will study a list of 5 words for 5 minutes, and those in the other group will study the same list for 7 minutes. Immediately after studying, the participants will read a list of 10 words and circle those that appeared on the original study list." Refine this experiment, giving specific reasons for any changes.

4. If you were investigating variables that affect helping behavior, would you be more likely to use a straightforward or staged manipulation? Why?

5. Design an experiment using a staged manipulation to test the hypothesis that when people are in a good mood, they are more likely to contribute to a charitable cause. Include a manipulation check in your design.

6. In a pilot study, Dr. Mori conducted a manipulation check and found no significant difference between the experimental conditions. Should she continue with the experiment? What should she do next? Explain your recommendations for Dr. Mori.