MODULE 12

Between-Participants Experimental Designs

LEARNING OBJECTIVES

- Explain a between-participants design.
- Differentiate independent variable and dependent variable.
- Differentiate control group and experimental group.
- Explain random assignment.
- Explain the relationship between confounds and internal validity.
- Describe the confounds of history, maturation, testing, regression to the mean, instrumentation, mortality, and diffusion of treatment.
- Explain what experimenter effects and participant effects are and how double-blind and single-blind experiments relate to these concepts.
- Differentiate floor and ceiling effects.
- Explain external validity.

Copyright 2010 Cengage Learning, All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part.

between-participants design: An experiment in which different participants are assigned to each group.

In a **between-participants design** the participants in each group are different, that is, different people serve in the control and experimental groups. The idea behind experimentation, as explained in Module 2, is that the researcher manipulates at least one variable (the *independent variable*) and measures at least one variable (the *dependent variable*). The independent variable has at least two groups or conditions. In other words, one of the most basic ideas behind an experiment is that there are at least two groups to compare. We typically refer to these two groups, or conditions, as the control group and the experimental group. The *control group* serves as the baseline, or "standard," condition. The *experimental group* receives some level of the independent variable. Although we begin by describing the two groups in an experiment as the experimental and control groups, an experiment may involve the use of two experimental groups with no control group. An experiment can also have more than two groups, that is, multiple experimental groups.

Experimentation requires control. We first have to control who is in the study. The sample must be representative of the population about whom we are trying to generalize. Ideally we accomplish this representation through the use of random sampling. We also need to control who participates in each condition, so we should use random assignment of participants to the two conditions. By randomly assigning participants, we are trying to make the two groups as equivalent as possible. In addition to controlling who serves in the study and in each condition, we need to control what happens during the experiment so that the only difference between conditions is between the levels of the independent variable. If, after controlling for all of these factors, we observe changes when the independent variable is manipulated, we can then conclude that the independent variable caused the changes in the dependent variable.

Let's revisit the example in Modules 9 and 10 on smoking and cancer in order to examine the difference between correlational research and experimental research. In those modules we said that there was a positive correlation between smoking and cancer. We also noted that no experimental evidence with humans supported a causal relationship between smoking and cancer. Why is this the case? Think about actually trying to design an experiment to determine whether smoking causes cancer in humans, keeping in mind the potential ethical problems with such an experiment.

Copyright 2010 Cengage Learning, All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part.

Let's first determine what the independent variable is. If you identified smoking behavior, you are correct. The control group would consist of people who do not smoke, and the experimental group would be the group who does smoke. To prevent confounding our study by previous smoking behavior, we would have to use only nonsmokers (those who had never smoked) in both the experimental and control groups. We would then randomly assign them to the smoking or nonsmoking groups. In addition to assigning participants to one of the two conditions, we would control all other aspects of their lives. This control means that all participants in the study must be treated exactly the same for the duration of the study except that half of them would smoke on a regular basis (we would decide when and how much) and half of them would not smoke at all. We would then determine the length of time the study should run. In this case participants would have to smoke for many years for us to assess any potential differences between groups. During this time all aspects of their lives that might contribute to cancer would have to be controlled, that is, held constant between the groups.

What would be the dependent variable? After several years had passed, we would begin to take measures on the two groups to determine whether there were any differences in cancer rates. Thus the cancer rate would be the dependent variable. If control was maximized and the experimental and control groups were treated exactly the same except for the level of the independent variable received, then any difference in cancer rate observed between the groups would have to be due to the only difference between them: the independent variable of smoking. This experimental study is illustrated in Figure 12.1.

You should be able to appreciate the problems associated with designing a true experiment to test whether smoking causes cancer in humans. First, it is not ethical for anyone to determine whether people should smoke or not. Second, it is not feasible to control all aspects of these individuals' lives for the period of time needed to conduct the study. It is for these reasons that there is no experimental study indicating that smoking causes cancer in humans.

It is perfectly feasible, however, to conduct experimental studies on other topics. For example, to study the effects of a mnemonic device (a study strategy) on memory, we could have one group use the device

Copyright 2010 Cengage Learning, All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part.

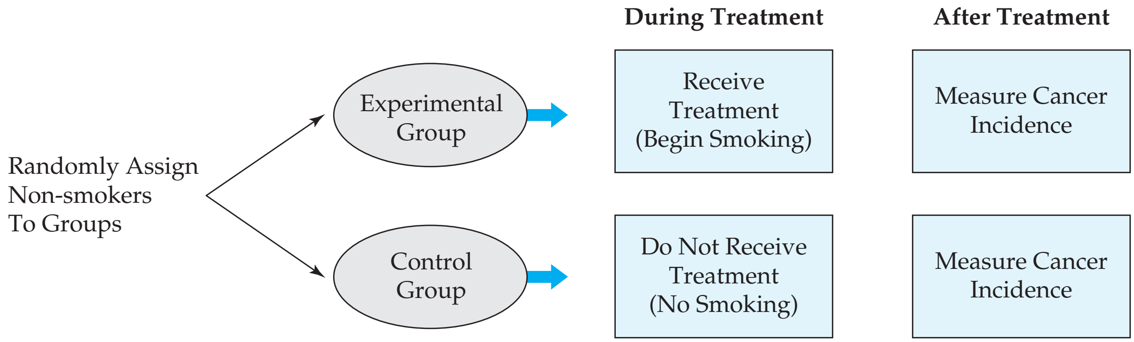

FIGURE **12.1** Experimental study of the effects of smoking on cancer rates

while studying and another group not use it. We could then give each participant a memory test and look for a difference in performance between the two groups. Assuming that everything else was held constant (controlled), any difference observed would have to be due to the independent variable. If the mnemonic-using group performed better, we could conclude that the mnemonic device caused memory to improve.

posttest-only control group design: An experimental design in which the dependent variable is measured after the manipulation of the independent variable.

This memory study is what is known as a simple **posttest-only control group design**. We start with a control group and an experimental group made up of equivalent participants. We administer the treatment (mnemonic or no mnemonic), and we take a posttest (after-treatment) measure. It is very important that the experimental and control groups are equivalent because we want to be able to conclude that any differences observed between the two groups are due to the independent variable, not to some other difference between them. We help to ensure equivalency of groups by using random assignment.

When we manipulate the independent variable, we must also ensure that the manipulation is valid, that is, that there really is a difference in the manner in which the two groups are treated. This determination appears fairly easy for the mnemonic device study; either the participants use the prescribed mnemonic device, or they do not. However, how do we actually know that those in the mnemonic group truly are using the device and that those in the control group are not using *any* type of mnemonic device? These are questions the researcher needs to address before beginning the study so that the instructions leave no doubt as to what the participants in each condition should be doing during the study.

Copyright 2010 Cengage Learning, All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part.

Finally, the researcher must measure the dependent variable (memory) to assess any effects of the independent variable. To be able to compare performance across the two groups, the same measurement device must be used for both groups. If the groups are equivalent at the beginning of the study and if the independent variable is adequately manipulated and the only difference between the two groups, then any differences observed on the dependent variable must be attributable to the independent variable.

pretest/posttest control group design: An experimental design in which the dependent variable is measured both before and after manipulation of the independent variable.

We could make the mnemonic device design slightly more sophisticated by using a **pretest/posttest control group design**, which involves adding a pretest to the design. This new design has the added advantage of ensuring that the participants are equivalent at the beginning of the study. This precaution is usually not considered necessary if participants are randomly assigned and if the researcher uses a sufficiently large sample of participants; as a general rule, having 20 to 30 participants per condition is considered adequate. There are disadvantages to pretest/posttest control group designs, including the possibility of increasing demand characteristics and experimenter effects (both discussed later in the module). The participants might guess before the posttest what is being measured in the study. If the participants make an assumption (either correct or incorrect) about the intent of the study, their behavior during the study may be changed from what would "normally" happen. With multiple testings there is also more opportunity for an experimenter to influence the participants. It is up to the researchers to decide which of these designs best suits their needs.

CONTROL AND CONFOUNDS

Obviously one of the most critical elements of an experiment is control, which must be maximized. If a researcher fails to control for something, then the study is open to confounds, that is, uncontrolled extraneous variables or flaws in an experiment (discussed in Module 10). If a study is confounded, then it is impossible to say whether changes in the dependent variable are caused by the independent variable or by the uncontrolled variable.

The problem for most psychologists is that maximizing control with human participants can be very difficult. In other disciplines control is not as difficult. For instance, marine biologists do not need to be as concerned about preexisting differences between sea snails because sea snails do not vary on as many dimensions as do humans (personality, intelligence, and rearing issues, among others, are not relevant as they are for humans). Because of the great variability among humans on all dimensions, psychologists need to be very concerned about

Copyright 2010 Cengage Learning, All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part.

preexisting differences. Consider the study on memory and mnemonic devices. A problem could occur if the differences in performance on the memory test resulted from the fact that based on chance, the more educated participants made up the experimental group and the less educated participants were in the control group. In this case we might have observed a difference between memory performance even if the experimental group had not used the mnemonic.

Even when we use random assignment as a means of minimizing differences between the experimental and control groups, we still need to think about control. So if conducting the study on memory and mnemonic devices, we should consider administering pretests as a means of assuring that the participants in the two groups are equivalent on any dimension (variable) that might affect memory performance. It is imperative that psychologists working with humans understand control and potential confounds due to human variability. If the basis of experimentation is that the control group and the experimental group (or the two experimental groups being compared) are as similar as possible except for differences in the independent variable, then it is up to the researcher to make sure that this situation is indeed the case. In short, the researcher needs to maximize the *internal validity* of the study, that is, the extent to which the results can be attributed to the manipulation of the independent variable rather than to some confounding variable. A study with good internal validity has no confounds and offers only one explanation for the results.

THREATS TO INTERNAL VALIDITY

There are several potential threats to the internal validity of a study. The confounds discussed below provide an overview of some potential problems and an opportunity to begin developing the critical thinking skills involved in designing a sound study. These confounds are most problematic for nonexperimental designs such as the quasi-experimental designs discussed in the previous chapter, but they may also pose a threat to experimental designs. Taking the precautions outlined here should indicate whether the confound is present in a study.

Nonequivalent Control Group

One of the most basic concerns in an experiment is that the participants in the control and experimental groups are equivalent at the beginning of the study. Using random sampling and random assignment is typically considered sufficient to address the potential problem of a nonequivalent control group. When random sampling and random assignment are not used, participant selection or assignment problems may result. In this case we would have a quasi-experimental design (discussed in Module 10), not a true experiment.

History

history effect: A threat to internal validity in which an outside event that is not a part of the manipulation of the experiment could be responsible for the results.

Changes in the dependent variable may be due to historical events that occur outside the study, leading to the confound known as a **history effect**. These events are most likely unrelated to the study but may nonetheless affect the dependent variable. Imagine that you are conducting a study on the effects of a certain program on stress reduction in college students. The study covers a 2-month period, during which students participate in your stress reduction program. If your posttest measure is taken during midterm or final exams, you might notice

Copyright 2010 Cengage Learning, All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part.

an increase in stress even though the participants are involved in a program intended to reduce stress. Not taking into account the historical point in the semester might lead you to an erroneous conclusion concerning the effectiveness of the stress reduction program. Notice also that a control group of equivalent participants would help reveal the confound in this study.

Maturation

maturation effect: A threat to internal validity in which participants' naturally occurring changes could be responsible for the observed results.

In research in which participants are studied over a period of time, a **maturation effect** can frequently be a problem. Participants mature physically, socially, and cognitively during the course of the study. Any changes in the dependent variable that occur over the course of the study therefore may be due to maturation and not to the independent variable of the study. Using a control group with equivalent participants indicates whether changes in the dependent variable are due to maturation; if they are, the participants in the control group change on the dependent variable during the course of the study even though they do not receive the treatment.

Testing

testing effect: A threat to internal validity in which repeated testing leads to better or worse scores.

In studies in which participants are measured numerous times, a **testing effect** may be a problem: repeated testing may lead to better or worse performance. Whereas many studies involve only pretest and posttest measures, others involve taking measures on an hourly, daily, weekly, or monthly basis. In these cases participants are exposed to the same or similar "tests" numerous times. As a result changes in performance on the test may be due to prior experience with it, not to the independent variable. As an example, if participants took the same math test before and after participating in a special math course, the improvement observed in scores might be due to participants' familiarity with and practice on the test items. This type of testing confound is sometimes referred to as a *practice effect*. Testing can also result in the opposite effect, a *fatigue effect* (sometimes referred to as a *negative practice effect*). Repeated testing fatigues the participants, and their performance declines as a result. Once again, having a control group of equivalent participants helps to control for testing confounds because researchers can see practice or fatigue effects in the control group.

Regression to the Mean

Statistical regression occurs when individuals are selected for a study because their scores on some measure were extreme, either extremely high or extremely low. *Regression to the mean* is a threat to internal validity in which extreme scores upon retesting tend to be less extreme, moving toward the mean. (To review this concept in more detail, refer to Module 10.)

Instrumentation

instrumentation effect: A threat to internal validity in which changes in the dependent variable may be due to changes in the measuring device.

An **instrumentation effect** occurs when the measuring device is faulty. Problems of consistency in measuring the dependent variable are most likely to occur when the measuring instrument is a human observer. The observer may become better at taking measures during the course of the study or may become fatigued. If the measures taken during the study are not taken consistently, then any change in the dependent variable may be due to measurement

Copyright 2010 Cengage Learning, All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part.

changes, not to the independent variable. Once again, having a control group of equivalent participants helps to identify this confound.

Mortality or Attrition

mortality (attrition): A threat to internal validity in which differential dropout rates may be observed in the experimental and control groups, leading to inequality between the groups.

Most research studies have a certain amount of **mortality**, or **attrition** (dropout). Most of the time the attrition is equal across experimental and control groups. It is of concern to researchers, however, when attrition is not equal across the groups. Assume that we begin a study with two equivalent groups of participants. If more participants leave one group than the other, then the two groups of participants are probably no longer equivalent, meaning that comparisons cannot be made between them.

What might cause differential attrition between groups? Imagine a study to test the effects of a program aimed at reducing smoking. We randomly select a group of smokers and then randomly assign half to the control group and half to the experimental group. The experimental group participates in the program to reduce smoking, but the heaviest smokers just cannot take its demands and quit. When we take a posttest measure on smoking, only the originally light to moderate smokers are left in the experimental group. Comparing them to the control group is pointless because the groups are no longer equivalent. Having a control group allows us to determine whether there is differential attrition across groups.

Diffusion of Treatment

diffusion of treatment: A threat to internal validity in which observed changes in the behaviors or responses of participants may be due to information received from other participants in the study.

When participants in a study are in close proximity to one another, a potential threat to internal validity is **diffusion of treatment**, that is, observed changes in the behaviors of participants may be due to information received from other participants. For instance, college students are frequently used as participants in research studies. Because many students live near one another and share classes, some may discuss an experiment in which they participated. If other students are planning to participate in the study in the future, the treatment has now been compromised because they know how some of the participants were treated. They know what is involved in one or more of the conditions in the study, and this knowledge may affect how they respond, regardless of the condition to which they are assigned. To control for this confound, researchers might try to run the participants in a study in large groups or within a short time span so that they do not have time to communicate with one another. In addition, researchers should stress to participants the importance of not discussing the experiment with anyone until it has ended.

Experimenter and Subject (Participant) Effects

experimenter effect: A threat to internal validity in which the experimenter, consciously or unconsciously, affects the results of the study.

Researchers invest considerable time and effort in designing experiments. Often this investment leads the researcher to consciously or unconsciously affect or bias the results of the study. Thus a researcher may unknowingly smile more when participants are behaving in the predicted manner and frown or grimace when participants are behaving in an undesirable manner. This type of **experimenter effect** is also referred to as *experimenter bias* because the results of the study are biased by the experimenter's expectations.

One of the most famous cases of experimenter effects is Clever Hans. Clever Hans was a horse that was purported to be able to do mathematical

Copyright 2010 Cengage Learning, All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part.

computations. Pfungst (1911) demonstrated that Hans's answers were based on experimenter effects. Hans supposedly solved mathematical problems by tapping out the answers with his hoof. A committee of experts who claimed Hans was receiving no cues from his questioners verified Hans's abilities. Pfungst later demonstrated that Hans in fact had no mathematical abilities and that tiny head and eye movements were Hans's signals to begin and end his tapping. When questioners asked Hans a question, they looked at Hans's hoof as he tapped out the answer. When Hans approached the correct number of taps, the questioners would unknowingly make a subtle head or eye movement in an upward direction. This movement was a cue to Hans to stop tapping.

single-blind experiment: An experimental procedure in which either the participants or the experimenter are blind to the manipulation being made.

double-blind experiment: An experimental procedure in which neither the experimenter nor the participant knows the condition to which each participant has been assigned; both parties are blind to the manipulation.

If a horse was clever enough to pick up on cues as subtle as these, imagine how human participants might respond to similar subtle cues provided by an experimenter. For this reason many researchers choose to combat experimenter effects by conducting blind experiments. There are two types of blind experiments: a single-blind experiment and a double-blind experiment. In a **single-blind experiment** either the experimenter or the participants are blind to the manipulation being made. The experimenter being blind in a single-blind experiment helps to combat experimenter effects. In a **double-blind experiment** neither the experimenter nor the participant knows the condition in which the participant is serving; both parties are blind. Obviously, the coordinator of the study has this information; however, the researcher responsible for interacting with the participants does not know and therefore cannot provide cues.

"IT WAS MORE OF A '<u>TRIPLE-BLIND</u>' TEST. THE PATIENTS DIDN'T KNOW WHICH ONES WERE GETTING THE REAL DRUG, THE DOCTORS DIDN'T KNOW, AND, I'M AFRAID, <u>NOBODY</u> KNEW."

© 2005 Sidney Harris, Reprinted with permission.

Copyright 2010 Cengage Learning, All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part.

Sometimes participants in a study bias the results based on their own expectations. They know they are being observed and hence may not behave naturally, or they may simply behave differently than when they are in more familiar situations. This type of confound is referred to as a **subject (participant) effect**. Sometimes subject effects are of a specific type. For instance, many participants try to be "good subjects," meaning that they try to determine what the researcher wants and to adjust their behavior accordingly. Such participants may be very sensitive to real or imagined cues from the researcher, referred to as *demand characteristics*. The subjects are trying to guess what characteristics the experimenter is in effect "demanding." Using either a single-blind experiment in which the participants are blind or a double-blind experiment helps to combat subject effects.

subject (participant) effect: A threat to internal validity in which the participant, consciously or unconsciously, affects the results of the study.

A special type of subject effect is often present in research on the effects of drugs and medical treatments. Most people report improvement when they are receiving a drug or other medical treatment. Some of this improvement may be caused by a *placebo effect*, that is, the improvement may be due not to the effects of the treatment but to the participant's expectation that the treatment will have an effect. For this reason drug and medical research must use a special placebo condition, or **placebo group**, a group of subjects who believe they are receiving treatment but in reality are not. Instead, they are given an inert pill or substance called a **placebo**. The placebo condition helps to distinguish between the actual effects of the drug and placebo effects. As an example, in a study on the effects of "ionized" wrist bracelets on musculoskeletal pain, researchers at the Mayo Clinic used a double-blind procedure in which half of the participants wore a so-called ionized bracelet and half of the participants wore a placebo bracelet. Both groups were told that they were wearing ionized bracelets intended to help with musculoskeletal pain. At the end of 4 weeks of treatment, both groups showed significant improvement in pain scores in comparison to baseline scores. No significant differences were observed between the groups. In other words, those wearing the placebo bracelet reported as much relief from pain as those wearing the ionized bracelet (Bratton et al., 2002).

placebo group: A group or condition in which participants believe they are receiving treatment but are not.

placebo: An inert substance that participants believe is a treatment.

Floor and Ceiling Effects

When conducting research, researchers must choose a measure for the dependent variable that is sensitive enough to detect differences between groups. If the measure is not sensitive enough, real differences may be missed. Although this confound does not result from an uncontrolled extraneous variable, it does represent a flaw in the experiment. For instance, measuring the weights of rats in pounds rather than ounces or grams is not advisable because no differences will be found. In this instance the insensitivity of the dependent variable is called a **floor effect**. All of the rats would be at the bottom of the measurement scale because the measurement scale is not sensitive enough to differentiate between such low scores. Similarly, attempting to weigh elephants on a bathroom scale would also lead to sensitivity problems; however, this is a **ceiling effect**. All of the elephants would weigh at the top of the scale (300 or 350 pounds, depending on the scale used), and any changes that might occur in weight as a result of the treatment variable would not be reflected in the

floor effect: A limitation of the measuring instrument that decreases its ability to differentiate between scores at the bottom of the scale.

ceiling effect: A limitation of the measuring instrument that decreases its ability to differentiate between scores at the top of the scale.

Copyright 2010 Cengage Learning, All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part.

dependent variable. A pretest can help to identify whether a measurement scale is sensitive enough. Participants should receive different scores on the dependent measure on the pretest. If all participants are scoring about the same (either very low or very high), then a floor or ceiling effect may be present.

IN REVIEW Threats to Internal Validity

Major Confounding Variables

Type of Confounding Variable	Description	Means of Controlling/ Minimizing
Nonequivalent control group	Problems in participant selection or assignment may lead to important differences between the participants assigned to the experimental and control groups	Use random sampling and random assignment of participants
History effect	Changes in the dependent variable may be due to outside events that take place during the course of the study	Use an equivalent control group
Maturation effect	Changes in the dependent variable may be due to participants maturing (growing older) during the course of the study	Use an equivalent control group
Testing effect	Changes in the dependent variable may be due to participants being tested repeatedly and getting either better or worse because of the repeated testings	Use an equivalent control group
Regression to the mean	Participants who are selected for a study because they are extreme (either high or low) on some variable may regress toward the mean and be less extreme at a later testing	Use an equivalent group of participants with extreme scores
Instrumentation effect	Changes in the dependent variable may be due to changes in the measuring device, either human or machine	Use an equivalent control group
Mortality (attrition)	Differential attrition or dropout in the experimental and control groups may lead to inequality between the groups	Monitor for differential loss of participants in experimental and control groups
Diffusion of treatment	Changes in the behaviors or responses of participants may be due to information they have received from others participating in the study	Attempt to minimize by testing participants all at once or as close together in time as possible
Experimenter and subject (participant) effects	Either experimenters or participants consciously or unconsciously affect the results of the study	Use a double-blind or single-blind procedure
Floor and ceiling effects	The measuring instrument used is not sensitive enough to detect differences	Ensure that the measuring instrument is reliable and valid before beginning the study

Copyright 2010 Cengage Learning, All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part.

CRITICAL THINKING CHECK 12.1

1. We discussed the history effect with respect to a study on stress reduction. Review that section and explain how having a control group of equivalent participants would help to reveal the confound of history.
2. Imagine that a husband and wife who are very tall (well above the mean for their respective height distributions) have a son. Would you expect the child to be as tall as his father? Why or why not?
3. While grading a large stack of essay exams, Professor Hyatt becomes tired and hence more lax in her grading standards. Which confound is relevant in this example? Why?

THREATS TO EXTERNAL VALIDITY

external validity: The extent to which the results of an experiment can be generalized.

In addition to internal validity, a study must have external validity for the results to be meaningful. **External validity** is the extent to which the results can be generalized beyond the participants used in the experiment and beyond the laboratory in which the experiment was conducted.

Generalization to Populations

college sophomore problem: An external validity problem that results from using mainly college sophomores as participants in research studies.

Generalization to the population being studied can be accomplished by randomly sampling participants from the population. Generalization to other populations, however, is problematic because most psychology research is conducted on college students, especially freshmen and sophomores; hardly a representative sample from the population at large. This problem—sometimes referred to as the **college sophomore problem** (Stanovich, 2007)—means that most conclusions are based on studies of young people with a late adolescent mentality who are still developing their own identities and attitudes (Cozby, 2001).

Does using college students as subjects in most research compromise research ideals? There are three responses to the college sophomore criticism (Stanovich, 2007). First, using college sophomores does not negate the findings of the study; it simply means that the study needs to be replicated with participants from other populations in order to aid in overcoming this problem. Second, in the research conducted in many areas of psychology such as sensory research, the college sophomore problem is not an issue. The auditory and visual systems of college sophomores function in the same manner as do those of the rest of the population. Third, the population of college students today is varied. They come from different socioeconomic backgrounds and geographic areas. They have varied family histories and educational experiences. Hence it is likely that college sophomores may be fairly representative of the general population.

Generalization from Laboratory Settings

Conducting research in a laboratory setting enables us to maximize control. We have discussed at several points the advantages of maximizing control, but control also has the potential disadvantage of creating an artificial

Copyright 2010 Cengage Learning, All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part.

environment. So we need to exercise caution when generalizing from the laboratory setting to the real world. This problem is often referred to in psychology as the *artificiality criticism* (Stanovich, 2007). Keep in mind, however, that the whole point of experimentation is to create a situation in which control is maximized in order to determine cause-and-effect relationships. Obviously we cannot relax our control in an experiment just to counter this criticism.

exact replication: Repeating a study using the same means of manipulating and measuring the variables as in the original study.

conceptual replication: A study based on another study that uses different methods, a different manipulation, or a different measure.

systematic replication: A study that varies from an original study in one systematic way—for example, by using a different number or type of participants, a different setting, or more levels of the independent variable.

How then can we address the artificiality criticism and the generalization issue? One way is through replication of the experiment so as to demonstrate that the result is reliable. A researcher might begin with an **exact replication**, that is, repeating the study in exactly the same manner. However, to more adequately address a problem such as the artificiality criticism, the researcher should consider a conceptual or systematic replication (Mitchell & Jolley, 2004). A **conceptual replication** tests the same concepts in a different way. Therefore we could use a different manipulation to assess its effect on the same dependent variable, or we could use the same manipulation and a different measure (dependent variable). A conceptual replication might also involve using other research methods to test the result. Accordingly, we might conduct an observational study (see Module 7) in addition to a true experiment to assess the generalizability of a finding. A **systematic replication** systematically changes one thing at a time and observes the effect, if any, on the results. For example, a study could be replicated with more or different participants, in a more realistic setting, or with more levels of the independent variable.

SUMMARY

Researchers should consider several factors when designing and evaluating a true experiment. First, they need to address the issues of control and possible confounds. The study needs to be designed with strong control and no confounds to maximize internal validity. Second, researchers should consider external validity in order to ensure that the study is as generalizable as possible while maintaining control. In addition, they should use the design most appropriate for the type of research they are conducting.

REVIEW OF KEY TERMS

between-participants design
posttest-only control group design
pretest/posttest control group design
history effect
maturation effect
testing effect
instrumentation effect
mortality (attrition)
diffusion of treatment
experimenter effect
single-blind experiment
double-blind experiment
subject (participant) effect
placebo group
placebo
floor effect
ceiling effect
external validity
college sophomore problem
exact replication
conceptual replication
systematic replication

Copyright 2010 Cengage Learning, All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part.