MODULE 10

Quasi-Experimental Designs

LEARNING OBJECTIVES

- Describe how quasi-experimental designs differ from correlational and experimental designs.
- Explain what a subject (participant) variable is.
- Differentiate single group designs and nonequivalent control group designs.
- Describe advantages and disadvantages of posttest-only designs and pretest/posttest designs.
- Explain a time-series design.

The term "quasi" (meaning "having some but not all of the features") preceding the term "experimental" indicates that we are dealing with a design that resembles an experiment but is not exactly an experiment. How does a quasi-experimental design differ from an experimental design? Sometimes the difference is the lack of a control group or a comparison group, that is, only one group is given a treatment and then assessed. At other times the independent variable is not a true manipulated independent variable; instead, it is a participant variable or a nonmanipulated independent variable. And finally, some designs may be considered quasi-experimental because participants were not randomly assigned to conditions, that is, they were already part of a group and the researcher attempted to manipulate a variable between preexisting groups.

NONMANIPULATED INDEPENDENT VARIABLES

In some quasi-experiments the researcher is interested in comparing groups of individuals (as is done in an experiment), but the groups occur naturally. In other words, participants are not assigned randomly to the groups. Notice the difference between this type of quasi-experimental design and correlational research. We are not simply looking for relationships between variables such as between smoking and cancer. In quasi-experimental research we are testing a hypothesis. An example is that individuals who have smoked for 20 years have a higher incidence of respiratory illness than nonsmokers. We would randomly select a group of individuals who had smoked for 20 years and a group of individuals who had never smoked to serve as a control. Thus rather than simply looking for a relationship between smoking and cancer or illness, we are comparing two groups to test a hypothesis.

nonmanipulated independent variable: The independent variable in a quasi-experimental design in which participants are not randomly assigned to conditions but rather come to the study as members of each condition.

The independent variable is referred to as a **nonmanipulated independent variable** because participants are not randomly assigned to the two groups. We are not truly manipulating smoking; participants come to the study as either smokers or nonsmokers. However, we do make comparisons between the groups. Consequently the study has the intent and "flavor" of an experiment without being a true experiment. Nonmanipulated independent variables are also known as *subject (participant) variables*. A subject variable, you may recall from Module 2, is a characteristic of the participant that cannot be changed such as ethnicity, gender, age, or political affiliation. If a study is designed to assess differences in individuals on some participant variable, by default it is a quasi-experiment and not a true experiment because it uses a nonmanipulated independent variable, that is, participants are not randomly assigned to conditions.

AN EXAMPLE: SNOW AND CHOLERA

In the 1850s in London, England, there were frequent outbreaks of cholera, an infection of the small intestine. The cause at the time was unknown, but the common theory was that cholera was somehow spread as people came in contact with cholera victims and shared or breathed the same air. This hypothesis was known as the effluvia theory. John Snow in his quest for the cause of cholera had an alternative hypothesis (Goldstein & Goldstein, 1978). Snow thought that people contracted cholera by drinking contaminated water. He based his hypothesis on the observation that of the several different water companies serving London, some provided water from upstream (it had not yet passed through the city and possibly become contaminated), whereas others used water from downstream (after it had passed through the city and possibly become contaminated).

To test this hypothesis, Snow used a quasi-experimental design. Obviously it was not feasible to use a true experimental design because it would have been impossible to randomly assign different houses to contract with a specific water company. Snow therefore had to look at houses that already received their water from a downstream company versus houses that received water from upstream. You should begin to see some of the problems inherent in quasi-experimental research. If people chose their water company, then there was most likely a reason for the choice. In most cases the reason was socioeconomic: The wealthy neighborhoods used upstream (more costly) companies, whereas the poor neighborhoods used downstream (less costly) companies. This socioeconomic distinction obviously presented a problem for Snow because he had no way of knowing whether differences in cholera incidence were due to the different water companies or to something else related to socioeconomic level such as diet, living conditions, or medical care.

Luckily for Snow, he was able to find one neighborhood in which socioeconomic status was stable but different houses received water from two different companies in an unsystematic manner. Hence the choice of water companies in this neighborhood appeared to be random. It was so random in fact that in some cases the choice of water company varied from house to house on a single street. Here was a naturally occurring situation in which socioeconomic level was controlled and the choice of water company varied. It was important, however, to ensure that not only the water company but also the contamination level of the water varied. Snow was lucky in this respect, too, because one company had moved upstream after a previous cholera epidemic, and the other company had stayed downstream. Snow calculated the number of deaths by cholera for individuals receiving water from upstream versus those receiving water from downstream. He found that there were 37 deaths per 10,000 households for the upstream company and 315 deaths per 10,000 households for the downstream company. Therefore it appeared that water contamination was responsible for the spread of cholera.

As a review the nonmanipulated independent variable in Snow's study was water company. This was a participant variable because individuals came to the study with their choice of water company already established. The dependent variable was the number of deaths by cholera. Snow observed

a difference in death rates between the two companies and concluded that the type of water (more contaminated versus less contaminated) appeared to be the cause. Snow was particularly lucky because of the naturally occurring situation in which socioeconomic level was controlled but water company varied. This type of control is often lacking in quasi-experimental research. Still, even with such control, there is not as much control as in an experiment because participants are not randomly assigned to conditions. Consequently it is still possible for uncontrolled differences between the groups to affect the outcome of the study.

IN REVIEW Quasi-Experimental Versus Correlational Methods

	Variables	Conclusions	Cautions
Correlational method	Two measured variables	The variables may be related in some way.	We cannot conclude that the relationship is causal.
Quasi-experimental method	Typically one nonmanipulated independent variable and one measured dependent variable	Systematic differences have been observed between two or more groups, but we cannot say that the nonmanipulated independent variable definitely caused the differences.	Due to confounds inherent in the use of nonmanipulated independent variables, there may be alternative explanations for the results.

CRITICAL THINKING CHECK 10.1

1. Which of the following variables would be a participant variable if used as a nonmanipulated independent variable in a quasi-experiment?

   gender — ethnicity
   religious affiliation — visual acuity
   amount of time spent studying — amount of alcohol consumed

2. How does the quasi-experimental method allow us to draw slightly stronger conclusions than the correlational method? Why is it that the conclusions drawn from quasi-experimental studies cannot be stated in as strong a manner as those from a true experiment?

TYPES OF QUASI-EXPERIMENTAL DESIGNS

The quasi-experimental design has several possible variations (Campbell & Stanley, 1963; Cook & Campbell, 1979; and Shadish, Cook, & Campbell, 2002). One distinction is whether there are one or two groups of participants. A second distinction has to do with how often measurements are taken. We begin by discussing quasi-experimental designs in which only one group of participants is observed. These designs include the single-group posttest-only design, the single-group pretest/posttest design, and the single-group time-series

design. We then consider designs that use two groups, which are referred to as *nonequivalent control group designs* and which include the nonequivalent control group posttest-only design, the nonequivalent control group pretest/posttest design, and the multiple-group time-series design.

Single-Group Posttest-Only Design

single-group posttest-only design: A design in which a single group of participants is given a treatment and then tested.

The **single-group posttest-only design** is the simplest quasi-experimental design. As the name implies, it involves the use of a single group of participants to whom some treatment is given. The participants are then assessed on the dependent variable. Research in education is frequently of this type. For example, a new educational technique—such as interactive learning, outcomes learning, or computer-assisted learning—is proposed, and school systems begin to adopt it. Posttest measures are then taken to determine the amount learned by students. However, there is neither a comparison group nor a comparison of the results to any previous measurements (usually because what is learned via the new method is so "different" from the old method that the claim is made that comparisons are not valid). This lack of comparison is the problem with this type of design: How can we claim a method is better when we cannot compare the results for the group who participated with the results for any other group or standard? This design is open to so many criticisms and potential flaws that results based on this type of study should always be interpreted with caution.

Single-group posttest-only designs are frequently reported in popular literature in which they are also frequently misinterpreted by those who read them. How many times have you read about people who lived through a certain experience or joined a particular group claiming that the experience or the group had an effect on their lives? These are examples of single-group posttest-only designs, and such designs cannot be used to draw conclusions about how an experience has affected the individuals involved. The change in their lives could be due to any number of variables other than the experience or the program.

Single-Group Pretest/Posttest Design

single-group pretest/posttest design: A design in which a single group of participants takes a pretest, then receives some treatment, and finally takes a posttest.

The **single-group pretest/posttest design** is an improvement over the posttest-only design in that measures are taken twice: before the treatment and after the treatment. The two measures can then be compared, and differences in the measures are assumed to be the result of the treatment. For instance, if a single group of depressed individuals wanted to receive treatment (counseling) for their depression, we would measure their level of depression before the treatment, we would then have them participate in the counseling, and finally, we would measure their level of depression after the treatment. Can you think of possible problems with this design? The greatest is the lack of a comparison group. With no comparison group, we do not know whether any observed change in depression is due to the treatment or to something else that may have happened during the time of the study. Maybe the pretest depression measure was taken right after the holidays when depression is higher than during the rest of the year for many people. Consequently the participants might have scored lower on the posttest depression measure regardless of the counseling.

Single-Group Time-Series Design

single-group time-series design: A design in which a single group of participants is measured repeatedly before and after a treatment.

The **single-group time-series design** involves using a single group of participants, taking multiple measures over a period of time before introducing the treatment, and then continuing to take several measures after the treatment. The advantage of this design is that the multiple measures allow us to see whether the behavior is stable before treatment and how, or if, it changes at the points in time at which measures are taken after treatment.

An oft-cited good example of a time-series design, discussed by Campbell (1969), was used to evaluate the 1955 crackdown on speeding in Connecticut. The state found it necessary to institute the crackdown after a record-high number of traffic fatalities occurred in 1955. A pretest/posttest design would simply have compared the number of fatalities before the crackdown with the number afterward. The number of deaths fell from 324 in 1955 to 284 in 1956. However, alternative hypotheses other than the crackdown could have been offered to explain the drop. Perhaps the number of deaths in 1955 was unusually high based on chance, that is, the number was just a "fluke." Campbell recommended a time-series design, examining traffic fatalities over an extended period. Figure 10.1 illustrates the results of this design, which includes traffic fatalities for the years 1951 through 1959. As can be seen in the figure, 1955 was a record-high year; after the crackdown the

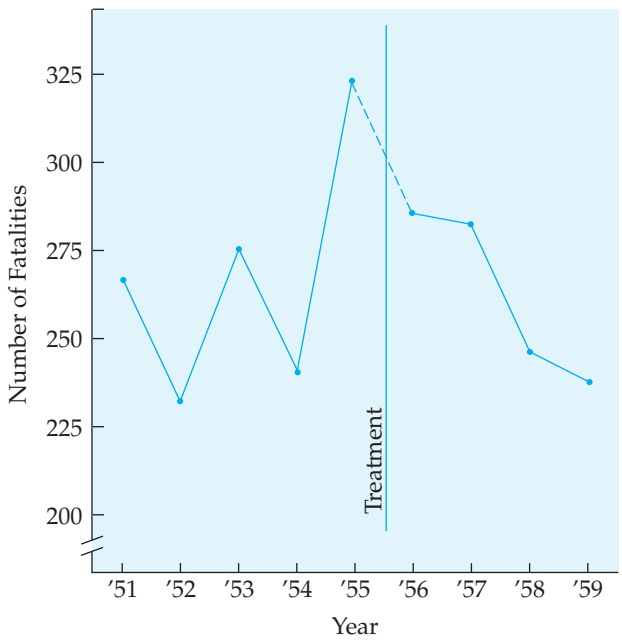

FIGURE **10.1** Connecticut traffic fatalities: 1951–1959

Source: D. T. Campbell, (1969). Reforms as experiments. *American Psychologist*, 24, 409–429. Copyright 1969 by the American Psychological Association. Reprinted with permission.

number of fatalities declined not only in 1956 but also in the 3 following years. Using the time-series design then allowed for a clearer interpretation than was possible with data from only 1955 and 1956.

Campbell still saw a problem with attributing the decline to the crackdown. The problem is statistical regression, or regression to the mean. Statistical regression occurs when individuals are selected for a study because their scores on some measure are extreme—either extremely high or extremely low. If we were studying students who scored in the top 10% on the SAT and we retested them on the SAT, we would expect them to do well again. Not all students, however, would score as well as they did originally because of *statistical regression*, often referred to as **regression to the mean**. 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. In other words, some of the students did well the first time due to chance or luck. What happens when they take the test a second time? They are not as lucky, and their scores regress toward the mean.

regression to the mean: A threat to internal validity in which extreme scores upon retesting tend to be less extreme, moving toward the mean.

Regression to the mean occurs in many situations other than in research studies. Many people think that a hex is associated with being on the cover of *Sports Illustrated* and that an athlete's performance declines after appearing on the cover. This decline can be explained by regression to the mean. Athletes are most likely to appear on the cover of *Sports Illustrated* after a very successful season or at the peak of their careers. What is most likely to happen after they have been performing exceptionally well over a period of time? They are likely to regress toward the mean and perform in a more average manner (Cozby, 2001). In a research study, having an equivalent control group of participants with extreme scores indicates whether changes in the dependent measure are due to regression to the mean or to the effects of the treatment variable.

Because of regression to the mean, with the very high death rate in 1955, we would expect a drop in the death rate for several years, whether there was a speeding crackdown or not, because the average death rate (calculated over several years) would remain the same. We will discuss Campbell's recommendation for an improved design shortly when we cover the multiple-group time-series design.

Nonequivalent Control Group Posttest-Only Design

nonequivalent control group posttest-only design: A design in which at least two nonequivalent groups are given a treatment and then a posttest measure.

The **nonequivalent control group posttest-only design** is similar to the single-group posttest-only design, but a nonequivalent control group is added as a comparison group. Notice that the control group is nonequivalent, meaning that participants are not assigned to either the experimental or the control group in a random manner. Instead, they are members of each group because of something that they chose or did, that is, they come to the study already a member of one of the groups. This design is similar to the quasi-experimental study conducted by Snow on cholera and discussed earlier in this module. Participants selected either the upstream or the downstream water company, and Snow took posttest measures on death rates by cholera. As noted earlier, Snow had some evidence that the two groups were somewhat equivalent on income level because they all lived in the same neighborhood. In many situations, however, there is no assurance that the two groups are at all equivalent on any

variable prior to the study. For this reason we cannot say definitively that the treatment is responsible for any observed changes in the groups. It could be that the groups were not equivalent at the beginning of the study, and hence the differences observed between the two groups on the dependent variable may be due to the nonequivalence of the groups and not to the treatment.

Nonequivalent Control Group Pretest/Posttest Design

nonequivalent control group pretest/posttest design: A design in which at least two nonequivalent groups are given a pretest, then a treatment, and finally a posttest.

An improvement over the previous design involves the addition of a pretest measure, making it a **nonequivalent control group pretest/posttest design**. This design is still not a true experimental one because as with the previous designs participants are not randomly assigned to the two conditions. However, a pretest allows us to assess whether the groups are equivalent on the dependent measure before the treatment is given to the experimental group. In addition, we can assess any changes that may have occurred in either group after treatment by comparing the pretest measures for each group with their posttest measures. Thus not only can we compare the performances of the two groups on both pretest and posttest measures, but we can compare performance within each group from the pretest to the posttest. If the treatment has some effect, then there should be a greater change from pretest to posttest for the experimental group than for the control group.

Williams (1986) and her colleagues used this design in a series of studies to assess the effects of television on communities. The researchers found a small Canadian town that had no television reception until 1973; they designated this town the Notel group. Life in Notel was then compared to life in two other communities: Unitel, which received only one station at the beginning of the study, and Multitel, which received four channels at the beginning of the study. A single channel was introduced to Notel at the beginning of the study. During the 2 years of the study Unitel began receiving three additional stations. The researchers measured such factors as participation in community activities and aggressive behavior in children in all three groups, both before and after the introduction of television in Notel. Results showed that after the introduction of television in Notel, there was a significant decline in participation in community activities and a significant increase in aggressive behavior in children.

Multiple-Group Time-Series Design

multiple-group time-series design: A design in which a series of measures are taken on two or more groups both before and after a treatment.

The logical extension of the previous design is to take more than one pretest and posttest. In a **multiple-group time-series design** several measures are taken on nonequivalent groups before and after treatment. Refer to the study of the crackdown on speeding in Connecticut following a high number of traffic fatalities in 1955. Converting that single-group time-series design to a multiple-group time-series design would involve finding a comparison group—a state that did not crack down on speeding—during the same time period. Campbell (1969) found four other states that did not crack down on speeding at the same time as Connecticut. Figure 10.2 presents the data from this design. As can be seen, the fatality rates in the states used as the control group remained fairly stable, while the fatality rates in Connecticut decreased. Based on these data, Campbell concluded that the crackdown had the desired effect on fatality rates.

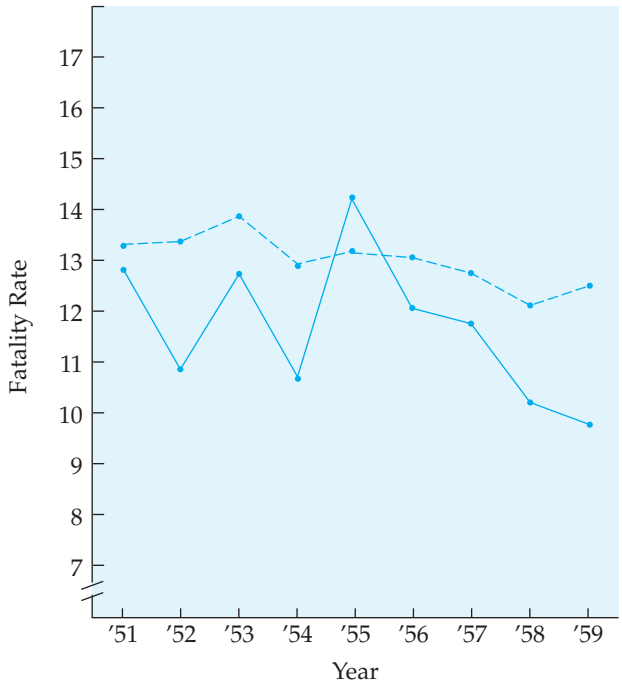

FIGURE **10.2** Multiple-group time-series design comparing Connecticut fatality rates (solid line) with the fatality rates of four other states (dashed line) used as a control group

Source: D. T. Campbell, (1969). Reforms as experiments. *American Psychologist*, 24, 409–429. Copyright 1969 by the American Psychological Association. Reprinted with permission.

INTERNAL VALIDITY AND CONFOUNDS IN QUASI-EXPERIMENTAL DESIGNS

confound: An uncontrolled extraneous variable or flaw in an experiment.

internal validity: The extent to which the results of an experiment can be attributed to the manipulation of the independent variable rather than to some confounding variable.

As we have pointed out several times, the results of quasi-experimental research need to be interpreted with caution because the design includes only one group or a nonequivalent control group. These results are always open to alternative explanations, or **confounds**, uncontrolled extraneous variables or flaws in an experiment. Because of the weaknesses in quasi-experimental designs, we can never conclude that the independent variable definitely caused any of the observed changes in the dependent variable. **Internal validity** is the extent to which the results of an experiment can be attributed to the manipulation of the independent variable rather than to some confounding variable. Quasi-experimental designs lack internal validity. We will continue to discuss internal validity and confounds when we cover true experimental designs in Module 12 as well as discussing how a true experiment helps to control for these confounds.

IN REVIEW Quasi-Experimental Designs

	Single Group Designs	Nonequivalent Control Group Designs
Posttest-only	Open to many confounds	Control group is nonequivalent
	No comparison group	No pretest measures to establish equivalence of groups
	No equivalent control group	Can compare groups on posttest measures, but differences may be due to treatment or confounds
Pretest/posttest	Can compare scores on pretest to those on posttest	Can compare between groups on pretest and posttest
	No equivalent control group for comparison	Can compare within groups from pretest to posttest
	If change is observed, it may be due to treatment or confounds	Because participants are not randomly assigned to groups, cannot say that they are equivalent
		If change is observed, may be due to treatment or confounds
Time series	Because many measures are taken, can see effect of treatment over time	Because many measures are taken, can see effect of treatment over time
	No control group for comparison	Nonequivalent control group available for comparison
	If change is observed, it may be due to treatment or confounds	Because participants are not randomly assigned to groups, cannot say that they are equivalent

CRITICAL THINKING CHECK 10.2

1. A researcher randomly selects a group of smokers and a group of nonsmokers and then measures lung disease in each group. What type of design is this? If the researcher observes a difference between the groups in the rate of lung disease, why can he or she not conclude that the difference is caused by smoking?
2. How are pretest/posttest designs an improvement over posttest-only designs?

SUMMARY

In this module you were introduced to quasi-experimental designs, a type of design that falls somewhere between a correlational design and a true experimental design. Important concepts related to quasi-experimental designs include nonmanipulated independent variables (participant variables), internal validity, and confounds. Quasi-experimental designs include both single-group designs and nonequivalent control group designs.