MODULE 9

Conducting Correlational Research

LEARNING OBJECTIVES

- Describe the difference between strong, moderate, and weak correlation coefficients.
- Draw and interpret scatterplots.
- Explain negative, positive, curvilinear, and no relationship between variables.
- Explain how assuming causality and directionality, the third-variable problem, restrictive ranges, and curvilinear relationships can be problematic when interpreting correlation coefficients.
- Explain how correlations allow us to make predictions.

When conducting correlational studies, researchers determine whether two naturally occurring variables (for example, height and weight or smoking and cancer) are related to each other. Such studies assess whether the variables are "co-related" in some way: Do tall people tend to weigh more than people of average height, or do those who smoke tend to have a higher-than-normal incidence of cancer? As we saw in Module 2, the correlational method is a type of nonexperimental method that describes the relationship between two measured variables. In addition to describing a relationship, correlations allow us to make predictions from one variable to another. If two variables are correlated, we can predict from one variable to the other with a certain degree of accuracy. Thus knowing that height and weight are correlated allows us to estimate, within a certain range, an individual's weight based on knowing the person's height.

Correlational studies are conducted for a variety of reasons. Sometimes it is impractical or ethically impossible to do an experimental study. For instance, it would be ethically impossible to manipulate smoking and assess whether it causes cancer in humans. How would you as a participant in an experiment like to be randomly assigned to the smoking condition and be told that you have to smoke a pack of cigarettes a day? Obviously this approach is not a viable experiment; however, one means of assessing the relationship between smoking and cancer is through correlational studies. In this type of study we can examine people who have already chosen to smoke and assess the degree of relationship between smoking and cancer.

Sometimes researchers choose to conduct correlational research because they are interested in measuring many variables and assessing the relationships between them. For example, they might measure various aspects of personality and assess the relationship between dimensions of personality.

MAGNITUDE, SCATTERPLOTS, AND TYPES OF RELATIONSHIPS

magnitude: An indication of the strength of the relationship between two variables.

Correlations vary in their **magnitude**, the strength of the relationship. Sometimes there is no relationship between variables, or the relationship may be weak; other relationships are moderate or strong. Correlations can also be represented graphically in a scatterplot or scattergram. In addition, relationships are of different types: positive, negative, none, or curvilinear.

Magnitude

The magnitude, or strength, of a relationship is determined by the correlation coefficient describing the relationship. As we saw in Module 6, a correlation coefficient is a measure of the degree of relationship between two variables; it can vary between −1.00 and +1.00. The stronger the relationship between the variables, the closer the coefficient is to either −1.00 or +1.00. The weaker the relationship between the variables, the closer the coefficient is to 0. You may recall from Module 6 that we typically discuss correlation coefficients as assessing a strong, moderate, or weak relationship, or no relationship at all. Table 9.1 provides general guidelines for assessing the magnitude of a relationship, but these ranges do not necessarily hold for all variables and all relationships.

A correlation coefficient of either −1.00 or +1.00 indicates a perfect correlation—the strongest relationship possible. For example, if height and weight were perfectly correlated (+1.00) in a group of 20 people, this coefficient would mean that the person with the highest weight was also the tallest person, the person with the second-highest weight was the second-tallest person, and so on down the line. In addition, in a perfect relationship each individual's score on one variable goes perfectly with his or her score on the other variable. For instance, this might mean that for every increase (decrease) in height of 1 inch, there is a corresponding increase (decrease) in weight of 10 pounds. If height and weight had a perfect negative correlation (−1.00), this coefficient would mean that the person with the highest weight was the shortest, the person with the second-highest weight was the second shortest, and so on, and that height and weight increased (decreased) by a set amount for each individual. It is very unlikely that you will ever observe a perfect correlation between two variables, but you may observe some very strong relationships between variables (±.70−.99). To sum up, whereas a correlation coefficient of ±1.00 represents a perfect relationship, a coefficient of 0 indicates no relationship between the variables.

Scatterplots

scatterplot: A figure that graphically represents the relationship between two variables.

A **scatterplot**, or scattergram, is a figure showing the relationship between two variables that graphically represents a correlation coefficient. Figure 9.1 presents a scatterplot of the height and weight relationship for 20 adults.

TABLE **9.1**

Estimates for Weak, Moderate, and Strong Correlation Coefficients

Correlation Coefficient	Strength of Relationship
±.70−1.00	Strong
±.30−.69	Moderate
±.00−.29	None (.00) to weak

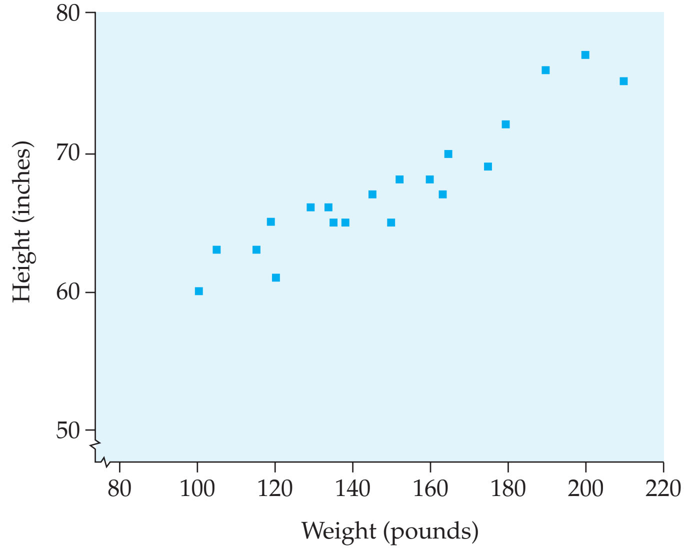


FIGURE **9.1** Scatterplot for height and weight

In a scatterplot two measurements are represented for each participant by the placement of a marker. In Figure 9.1 the horizontal *x*-axis shows the participant's weight, and the vertical *y*-axis shows height. The two variables could be reversed on the axes, and it would make no difference in the scatterplot. This scatterplot shows an upward trend, and the points cluster in a linear fashion. The stronger the correlation is, the more tightly the data points cluster around an imaginary line through their center. When there is a perfect correlation (± 1.00), the data points all fall on a straight line. In general, a scatterplot may show four basic patterns: a positive relationship, a negative relationship, no relationship, or a curvilinear relationship.

Positive Relationships

The relationship represented in Figure 9.2a shows a positive correlation, one in which there is a direct relationship between the two variables: An increase in one variable is related to an increase in the other, and a decrease in one is related to a decrease in the other. Notice that this scatterplot is similar to the one in Figure 9.1. The majority of the data points fall along an upward angle (from the lower left corner to the upper right corner). In this example a person who scored low on one variable also scored low on the other, an individual with a mediocre score on one variable had a mediocre score on the other, and anyone who scored high on one variable also scored high on the other. In other words, an increase (decrease) in one variable is accompanied by an increase (decrease) in the other; as variable *x* increases (or decreases), variable *y* does the same. If the data in Figure 9.2a represented height and weight measurements, we could say that those who are taller tend to weigh more, whereas

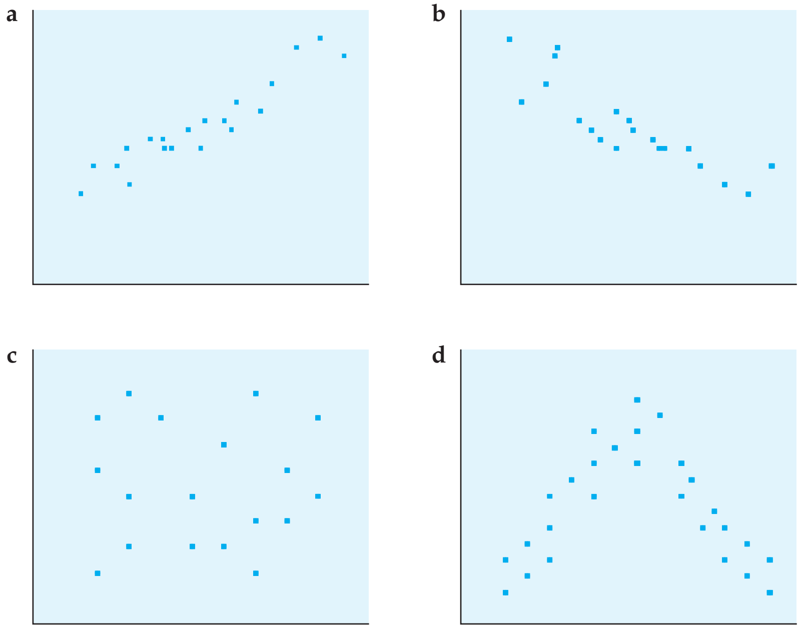


FIGURE **9.2** Possible types of Correlational relationships: (a) positive; (b) negative; (c) none; (d) curvilinear

those who are shorter tend to weigh less. Notice also that the relationship is linear: We could draw a straight line representing the relationship between the variables, and the data points would all fall fairly close to that line.

Negative Relationships

Figure 9.2b represents a negative relationship between two variables. Notice that in this scatterplot the data points extend from the upper left to the lower right. This negative correlation indicates that an increase in one variable is accompanied by a *decrease* in the other variable. This correlation represents an inverse relationship: The more of variable x that we have, the less we have of variable y. Assume that this scatterplot represents the relationship between age and eyesight. As age increases, the ability to see clearly tends to decrease—a negative relationship.

No Relationship

As shown in Figure 9.2c, it is also possible to observe no meaningful relationship between two variables. In this scatterplot the data points are scattered randomly. As you would expect, the correlation coefficient for these data is very close to 0 (−.09).

Curvilinear Relationships

A correlation coefficient of 0 indicates no meaningful relationship between two variables. However, it is also possible for a correlation coefficient of 0 to indicate a curvilinear relationship, as illustrated in Figure 9.2d. Imagine that this graph represents the relationship between psychological arousal (the x-axis) and performance (the y-axis). Individuals perform better when they are moderately aroused than when arousal is either very low or very high. The correlation coefficient for these data is also very close to 0 ($-.05$). Think about why this strong curvilinear relationship leads to a correlation coefficient close to 0. The strong positive relationship depicted in the left half of the graph essentially cancels out the strong negative relationship in the right half of the graph. Although the correlation coefficient is very low, we would not conclude that there is no relationship between the two variables. As the figure shows, the variables are very strongly related to each other in a curvilinear manner, with the points being tightly clustered in an inverted U shape.

Correlation coefficients only tell us about linear relationships. Thus even though there is a strong relationship between the two variables in Figure 9.2d, the correlation coefficient does not indicate this relationship because it is curvilinear. For this reason it is important to examine a scatterplot of the data in addition to calculating a correlation coefficient. Alternative statistics (beyond the scope of this text) can be used to assess the degree of curvilinear relationship between two variables.

IN REVIEW Relationships Between Variables

	Type of Relationships			
	Positive	**Negative**	**None**	**Curvilinear**
Description of Relationship	Variables increase and decrease together	As one variable increases, the other decreases in an inverse relationship	Variables are unrelated and do not move together in any way	Variables increase together up to a point and then as one continues to increase, the other decreases
Description of scatterplot	Data points are clustered in a linear pattern extending from lower left to upper right	Data points are clustered in a linear pattern extending from upper left to lower right	There is no pattern to the data points they are scattered all over the graph	Data points are clustered in a curved linear pattern forming a U shape or an inverted U shape
Example of variables related in this manner	Smoking and cancer	Mountain elevation and temperature	Intelligence and weight	Memory and age

CRITICAL THINKING CHECK 9.1

1. Which of the following correlation coefficients represents the weakest relationship between two variables?
   $-.59$ $+.10$ -1.00 $+.76$
2. Explain why a correlation coefficient of 0 or close to 0 may not mean that there is no relationship between the variables.
3. Draw a scatterplot representing a strong negative correlation between depression and self-esteem. Make sure you label the axes correctly.

MISINTERPRETING CORRELATIONS

Correlational data are frequently misinterpreted, especially when presented by newspaper reporters, talk show hosts, and television newscasters. Here we discuss some of the most common problems in interpreting correlations. Remember, a correlation simply indicates that there is a weak, moderate, or strong relationship (either positive or negative) or no relationship between two variables.

The Assumptions of Causality and Directionality

The most common error made when interpreting correlations is assuming that the relationship observed is causal in nature: that a change in variable A *causes* a change in variable B. Correlations simply identify relationships; they do not indicate causality. For example, a commercial recently appeared on television sponsored by an organization promoting literacy. The statement was made at the beginning of the commercial that a strong positive correlation had been observed between illiteracy and drug use in high school students (those high on the illiteracy variable also tended to be high on the drug use variable). The commercial concluded with a statement along the lines of "Let's stop drug use in high school students by making sure they can all read." Can you see the flaw in this conclusion? The commercial did not air for very long, probably because someone pointed out the error.

causality: The assumption that a correlation indicates a causal relationship between two variables.

directionality: The inference made with respect to the direction of a causal relationship between two variables.

This commercial made the twin errors of assuming causality and directionality. **Causality** refers to the assumption that the correlation between two variables indicates a causal relationship, and **directionality** refers to the inference made with respect to the direction of a causal relationship between two variables. The commercial assumed that illiteracy was causing drug use; it claimed that if illiteracy were lowered, then drug use would also be lowered. As we know, a correlation between two variables indicates only that they are related, that is, they vary together. Although it is possible that one variable causes changes in the other, we cannot draw this conclusion from correlational data.

Research on smoking and cancer illustrates this limitation of correlational data. For research with humans we have only correlational data indicating a positive correlation between smoking and cancer. Because the data are correlational, we cannot conclude that there is a causal relationship. In this situation it is probable that the relationship is causal. However, based solely on correlational data, we cannot draw that conclusion, nor can we assume the direction of the relationship. Thus the tobacco industry could argue that, yes, there is a correlation between smoking and cancer, but maybe cancer causes smoking, or maybe individuals predisposed to cancer are more attracted to smoking cigarettes. Even though experimental data based on research with laboratory animals indicate that smoking causes cancer, the tobacco industry questions whether the research is applicable to humans and for years continued to state that no research had produced evidence of a causal link between smoking and cancer in humans.

A classic example of the assumption of causality and directionality with correlational data occurred when researchers observed a strong negative correlation between eye movement patterns and reading ability in children. Poor readers tended to make more erratic eye movements than normal, more movements from right to left, and more stops per line of text. Based on this correlation, some researchers assumed causality and directionality: They presumed that poor oculomotor skills caused poor reading and proposed programs for "eye movement training." Many elementary school students who were poor readers spent time in such training, supposedly developing oculomotor skills in the hope that these skills would improve their reading ability. Experimental research later provided evidence that the relationship between eye movement patterns and reading ability is indeed causal, but that the direction of the relationship is the reverse: poor reading causes more erratic eye movements! Children who are having trouble reading need to go back over the information more and stop and think about it more. When children improve their reading skills (i.e., improve recognition and comprehension), their eye movements become smoother (Olson & Forsberg, 1993). Because of the errors of assuming causality and directionality, many children never received the appropriate training to improve their reading ability.

The Third-Variable Problem

When we interpret a correlation, it is important to remember that although the correlation between the variables may be very strong, the relationship may be the result of a third variable that influences both of the measured variables. The **third-variable problem** results when a correlation between two variables is dependent on another (third) variable.

third-variable problem: The problem of a correlation between two variables being dependent on another (third) variable.

A good example of the third-variable problem is a well-cited study conducted by social scientists and physicians in Taiwan (Li, 1975). The researchers attempted to identify the variables that best predicted the use of birth control; a question of interest to the researchers because of overpopulation problems in Taiwan. They collected data on various behavioral and environmental variables and found that the variable most strongly correlated with contraceptive use was the number of electrical appliances (yes, electrical appliances—stereos, toasters, televisions, and so on) in the home. If we take this correlation at face value, it means that individuals who use many electrical appliances tend also to use contraceptives, whereas those with fewer electrical appliances tend to use contraceptives less.

It should be obvious that this relationship is not causal (buying electrical appliances does not cause individuals to use birth control, nor does using birth control cause individuals to buy electrical appliances). Thus we probably do not have to worry about people assuming either causality or directionality when interpreting this correlation. The problem is a third variable. In other words, the relationship between electrical appliances and contraceptive use is not really a meaningful relationship; other variables are tying them together. Can you think of other ways in which individuals who use contraceptives and who have a large number of appliances might be similar? Education is a possible third variable. Individuals with a higher education level tend to be better informed about contraceptives and also tend to have a higher socioeconomic status (they get better paying jobs). Their higher socioeconomic status allows them to buy more "things," including electrical appliances.

partial correlation: A correlational technique that involves measuring three variables and then statistically removing the effect of the third variable from the correlation of the remaining two.

It is possible statistically to determine the effects of a third variable by using a correlational procedure known as **partial correlation,** which involves measuring all three variables and then statistically removing the effect of the third variable from the correlation of the remaining two. If the third variable (in this case, education) is responsible for the relationship between electrical appliances and contraceptive use, then the correlation should disappear when the effect of education is removed, or partialed out.

Restrictive Range

restrictive range: A variable that is truncated and has limited variability.

The idea behind measuring a correlation is that we assess the degree of relationship between two variables. Variables by definition must vary. When a variable is truncated, we say that it has a **restrictive range,** that is, the variable does not vary enough. Look at Figure 9.3a, which represents a scatterplot of SAT scores and college GPAs for a group of students. SAT scores and GPAs are positively correlated. Neither of these variables is restricted in range (for this group of students, SAT scores vary from 400 to 1600 and GPAs vary from 1.5 to 4.0), so we have the opportunity to observe a relationship between the variables. Now look at Figure 9.3b, which represents the correlation between the same two variables, except the range on the SAT variable is restricted to those who scored between 1000 and 1150. The SAT variable has been restricted, or truncated, and does not "vary" very much. As a result the opportunity to observe a correlation has been diminished. Even if there were a strong relationship between these variables, we could not observe it because of the restricted range of one of the variables. Thus when interpreting and using correlations, beware of variables with restricted ranges.

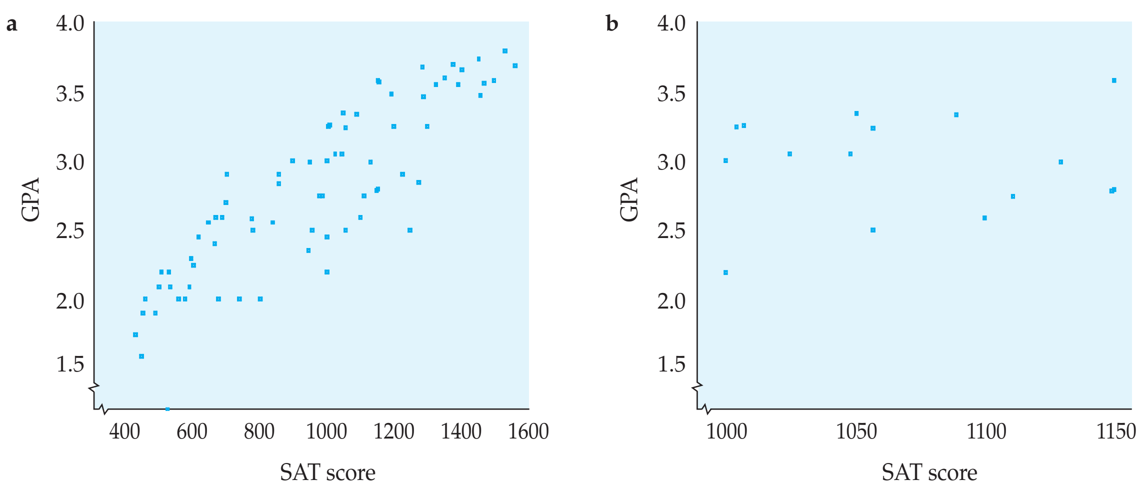


FIGURE **9.3** Restrictive range and correlation

Curvilinear Relationships

Curvilinear relationships and the caution in interpreting them were discussed earlier in the module. Because correlations are a measure of linear relationships, when a relationship is curvilinear, a correlation coefficient does not adequately indicate the degree of relationship between the variables. If necessary, look back over the previous section on curvilinear relationships in order to refresh your memory concerning them.

IN REVIEW Misinterpreting Correlations

	Types of Misinterpretations			
	Causality and Directionality	**Third Variable**	**Restrictive Range**	**Curvilinear Relationship**
Description of Misinterpretation	We assume that the correlation is causal and that one variable causes changes in the other.	Other variables are responsible for the observed correlation.	One or more of the variables is truncated or restricted, and the opportunity to observe a relationship is minimized.	The curved nature of the relationship decreases the observed correlation coefficient.
Examples	We assume that smoking causes cancer or that illiteracy causes drug abuse because a correlation has been observed.	We find a strong positive relationship between birth control and the number of electrical appliances.	If SAT scores are restricted (limited in range), the correlation between SAT and GPA appears to decrease.	As arousal increases, performance increases up to a point; as arousal continues to increase, performance decreases.

CRITICAL THINKING CHECK 9.2

1. "I have recently observed a strong negative correlation between depression and self-esteem." Explain what this statement means. Make sure you avoid the misinterpretations described in the text.
2. General State University officials recently investigated the relationship between SAT scores and GPAs (at graduation) for its senior class. They were surprised to find a weak correlation between these two variables. They know they have a grade inflation problem (the whole senior class graduated with GPAs of 3.0 or higher), but they are unsure how this might help account for the low correlation observed. Can you explain?

Prediction and Correlation

Correlation coefficients not only describe the relationship between variables, but they also allow us to make predictions from one variable to another. Correlations between variables indicate that when one variable is present at a certain level, the other also tends to be present at a certain level. Notice the wording. The statement is qualified by the phrase "tends to." We are not saying that a prediction is guaranteed or that the relationship is causal but simply that the variables seem to occur together at specific levels. Think about some of the examples used in this module. Height and weight are positively correlated. One is not causing the other; nor can we predict an individual's weight exactly based on height (or vice versa). But because the two variables are correlated, we can predict with a certain degree of accuracy what an individual's approximate weight might be if we know the person's height.

Let's take another example. We have noted a correlation between SAT scores and college freshman GPAs. Think about the purpose of the SAT. College admissions committees use the test as part of the admissions procedure because there is a positive correlation between SAT scores and college freshman GPAs. Individuals who score high on the SAT tend to have higher college freshman GPAs; those who score lower on the SAT tend to have lower college freshman GPAs. Therefore knowing students' SAT scores can help predict, with a certain degree of accuracy, their freshman GPAs and their potential for success in college. At this point some of you are probably saying, "But that isn't true for me. I scored poorly (or very well) on the SAT, and my GPA is great (or not so good)." Statistics tell us only the trend for most people in the population or sample. There are always outliers—the few individuals who do not fit the trend. Most people, however, are going to fit the pattern.

Think about another example. There is a strong positive correlation between smoking and cancer, but you may know someone who has smoked for 30 or 40 years and does not have cancer or any other health problems. Does this one individual negate the fact that there is a strong relationship between smoking and cancer? No. To claim that it does would be a classic **person-who argument**, that is, arguing that a well established statistical trend is invalid because we know a "person who" went against the trend (Stanovich, 2007). A counterexample does not change the existence of a strong statistical relationship between the variables nor that you are increasing your chance of getting cancer if you smoke. Because of the correlation between the variables, we can predict (with a fairly high degree of accuracy) who might get cancer based on knowing a person's smoking history.

person-who argument: Arguing that a well-established statistical trend is invalid because we know a "person who" went against the trend.

SUMMARY

After reading this module, you should have an understanding of the correlational research method, which allows researchers to observe relationships between variables, and of correlation coefficients, the statistics that assess the relationship. Correlations vary in type (positive, negative, none, or curvilinear) and magnitude (weak, moderate, or strong). The pictorial representation of a correlation is a scatterplot. A scatterplot allows us to see the relationship, facilitating its interpretation.

Several errors are commonly made when interpreting correlations, including assuming causality and directionality, overlooking a third variable, having a restrictive range on one or both variables, and assessing a curvilinear relationship. Knowing that two variables are correlated allows researchers to make predictions from one variable to the other.

REVIEW OF KEY TERMS

magnitude	causality	third-variable problem	restrictive range
scatterplot	directionality	partial correlation	person-who argument

MODULE EXERCISES

(Answers to odd-numbered exercises appear in Appendix A.)

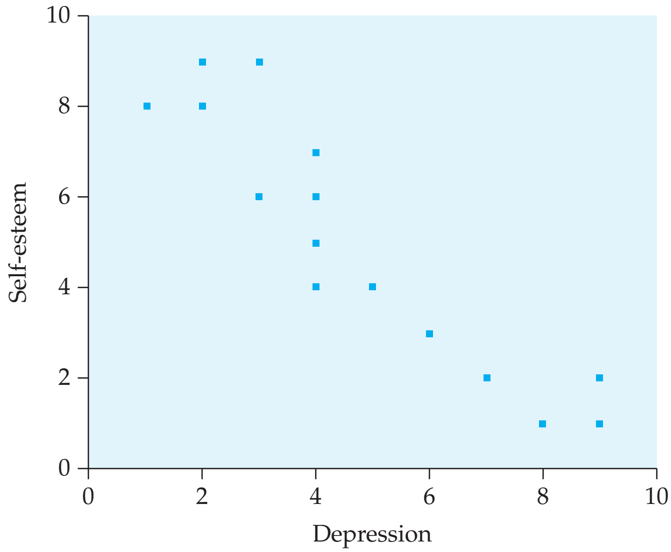
1. A health club recently conducted a study of its members and found a positive relationship between exercise and health. It was claimed that the correlation coefficient between the variables of exercise and health was +1.25. What is wrong with this statement? In addition, it was stated that this finding proved that an increase in exercise increases health. What is wrong with this statement?
2. Draw a scatterplot indicating a strong negative relationship between the variables of income and mental illness. Be sure to label the axes correctly.
3. We have mentioned several times that there is a fairly strong positive correlation between SAT scores and freshman GPAs. The admissions process for graduate school is based on a similar test, the GRE, which like the SAT has a total point range of 400 to 1,600. Let's assume that graduate schools do not accept anyone who scores below 1,000 and that a GPA below 3.00 represents failing work in graduate school. What would we expect the correlation between GRE scores and graduate school GPAs to be like in comparison to the correlation between SAT scores and college GPAs? Why would we expect this?

CRITICAL THINKING CHECK ANSWERS

9.1

1. +.10
2. A correlation coefficient of 0 or close to 0 may indicate no relationship or a weak relationship. However, if the relationship is curvilinear, the correlation coefficient could also be 0 or close to this. In this latter case there is a relationship between the two variables, but because the relationship is curvilinear, the correlation coefficient does not truly represent the strength of the relationship.

3.



9.2

1. A strong negative correlation between depression and self-esteem means that as individuals become more depressed, their self-esteem tends to decrease, whereas when individuals become less depressed, their self-esteem tends to increase. It does not mean that one variable causes changes in the other but simply that the variables tend to move together in a certain manner.

2. General State University officials observed such a weak correlation between GPAs and SAT scores because of a restrictive range on the GPA variable. Because of grade inflation, the whole senior class graduated with a GPA of 3.0 or higher. This restriction on one of the variables lessens the opportunity to observe a correlation.

WEB RESOURCES

Check your knowledge of the content and key terms in this module with a practice quiz and interactive flashcards at **www.cengage.com/psychology/jackson**, or for step-by-step practice and information, check out the Statistics and Research Methods Workshops at **www.cengage.com/psychology/workshops**.

LAB RESOURCES

For hands-on experience using the research methods described in this module, see Chapter 3 ("Correlation Research") in *Research Methods Laboratory Manual for Psychology*, 2nd ed., by William Langston (Belmont, CA: Wadsworth, 2005).

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

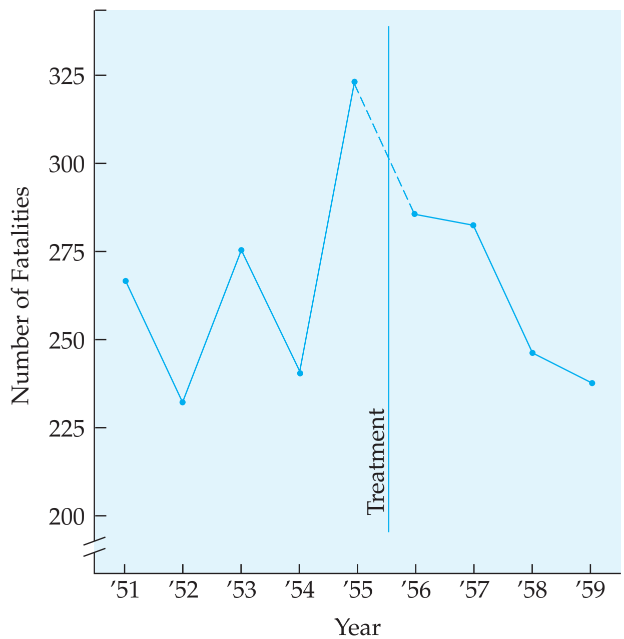


FIGURE **10.1** Connecticut traffic fatalities: 1951–1959

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.

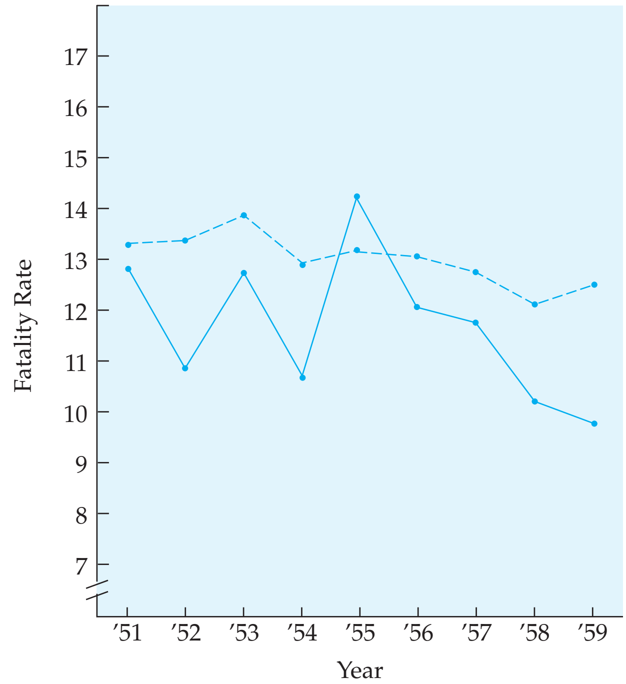


FIGURE **10.2** Multiple-group time-series design comparing Connecticut fatality rates (solid line) with the fatality rates of four other states (dashed line) used as a control group

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.