PART VI

NON-EXPERIMENTAL RESEARCH

What do the following classic studies have in common?

- Stanley Milgram found that about two thirds of his research participants were willing to administer dangerous shocks to another person just because they were told to by an authority figure (Milgram, 1963)[1].
- Elizabeth Loftus and Jacqueline Pickrell showed that it is relatively easy to "implant" false memories in people by repeatedly asking them about childhood events that did not actually happen to them (Loftus & Pickrell, 1995)[2].
- John Cacioppo and Richard Petty evaluated the validity of their Need for Cognition Scale—a measure of the extent to which people like and value thinking—by comparing the scores of university professors with those of factory workers (Cacioppo & Petty, 1982)[3].
- David Rosenhan found that confederates who went to psychiatric hospitals claiming to have heard voices saying things like "empty" and "thud" were labeled as schizophrenic by the hospital staff and kept there even though they behaved normally in all other ways (Rosenhan, 1973)[4].

The answer for purposes of this chapter is that they are not experiments. In this chapter, we look more closely at non-experimental research. We begin with a general definition of non-experimental research, along with a discussion of when and why non-experimental research is more appropriate than experimental research. We then look separately at two important types of non-experimental research: correlational research and observational research.

1. Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, 67, 371–378.
2. Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals*, 25, 720–725.
3. Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116–131.
4. Rosenhan, D. L. (1973). On being sane in insane places. *Science*, 179, 250–258.

28. Overview of Non-Experimental Research

Learning Objectives

1. Define non-experimental research, distinguish it clearly from experimental research, and give several examples.
2. Explain when a researcher might choose to conduct non-experimental research as opposed to experimental research.

What Is Non-Experimental Research?

Non-experimental research is research that lacks the manipulation of an independent variable. Rather than manipulating an independent variable, researchers conducting non-experimental research simply measure variables as they naturally occur (in the lab or real world).

Most researchers in psychology consider the distinction between experimental and non-experimental research to be an extremely important one. This is because although experimental research can provide strong evidence that changes in an independent variable cause differences in a dependent variable, non-experimental research generally cannot. As we will see, however, this inability to make causal conclusions does not mean that non-experimental research is less important than experimental research. It is simply used in cases where experimental research is not able to be carried out.

When to Use Non-Experimental Research

As we saw in the last chapter, experimental research is appropriate when the researcher has a specific research question or hypothesis about a causal relationship between two variables—and it is possible, feasible, and ethical to manipulate the independent variable. It stands to reason, therefore, that non-experimental research is appropriate—even necessary—when these conditions are not met. There are many times in which non-experimental research is preferred, including when:

- the research question or hypothesis relates to a single variable rather than a statistical relationship between two variables (e.g., how accurate are people's first impressions?).
- the research question pertains to a non-causal statistical relationship between variables (e.g., is there a correlation between verbal intelligence and mathematical intelligence?).
- the research question is about a causal relationship, but the independent variable cannot be manipulated or participants cannot be randomly assigned to conditions or orders of conditions for practical or ethical reasons (e.g., does damage to a person's hippocampus impair the formation of long-term memory traces?).
- the research question is broad and exploratory, or is about what it is like to have a particular experience (e.g., what is it like to be a working mother diagnosed with depression?).

Again, the choice between the experimental and non-experimental approaches is generally dictated by the nature of the research question. Recall the three goals of science are to describe, to predict, and to explain. If the goal is to explain and the research question pertains to causal relationships, then the experimental approach is typically preferred. If the goal is to describe or to predict, a non-experimental approach is appropriate. But the two approaches can also be used to address the same research question in complementary ways. For example, in Milgram's original (non-experimental) obedience study, he was primarily interested in one variable—the extent to which participants obeyed the researcher when he told them to shock the confederate—and he observed all participants performing the same task under the same conditions. However, Milgram subsequently conducted experiments to explore the factors that affect obedience. He manipulated several independent variables, such as the distance between the experimenter and the participant, the participant and the confederate, and the location of the study (Milgram, 1974)[1].

Types of Non-Experimental Research

Non-experimental research falls into two broad categories: correlational research and observational research.

The most common type of non-experimental research conducted in psychology is correlational research. Correlational research is considered non-experimental because it focuses on the statistical relationship between two variables but does not include the manipulation of an independent variable. More specifically, in **correlational research**, the researcher measures two variables with little or no attempt to control extraneous variables and then assesses the relationship between them. As an example, a researcher interested in the relationship between self-esteem and school achievement could collect data on students' self-esteem and their GPAs to see if the two variables are statistically related.

Observational research is non-experimental because it focuses on making observations of behavior in a natural or laboratory setting without manipulating anything. Milgram's original obedience study was non-experimental in this way. He was primarily interested in the extent to which participants obeyed the researcher when he told them to shock the confederate and he observed all participants performing the same task under the same conditions. The study by Loftus and Pickrell described at the beginning of this chapter is also a good example of observational research. The variable was whether participants "remembered" having experienced mildly traumatic childhood events (e.g., getting lost in a shopping mall) that they had not actually experienced but that the researchers asked them about repeatedly. In this particular study, nearly a third of the participants "remembered" at least one event. (As with Milgram's original study, this study inspired several later experiments on the factors that affect false memories).

Cross-Sectional, Longitudinal, and Cross-Sequential Studies

When psychologists wish to study change over time (for example, when developmental psychologists wish to study aging) they usually take one of three non-experimental approaches: cross-sectional,

1. Milgram, S. (1974). *Obedience to authority: An experimental view*. New York, NY: Harper & Row.

longitudinal, or cross-sequential. **Cross-sectional studies** involve comparing two or more pre-existing groups of people (e.g., children at different stages of development). What makes this approach non-experimental is that there is no manipulation of an independent variable and no random assignment of participants to groups. Using this design, developmental psychologists compare groups of people of different ages (e.g., young adults spanning from 18-25 years of age versus older adults spanning 60-75 years of age) on various dependent variables (e.g., memory, depression, life satisfaction). Of course, the primary limitation of using this design to study the effects of aging is that differences between the groups other than age may account for differences in the dependent variable. For instance, differences between the groups may reflect the generation that people come from (a **cohort effect**) rather than a direct effect of age. For this reason, **longitudinal studies**, in which one group of people is followed over time as they age, offer a superior means of studying the effects of aging. However, longitudinal studies are by definition more time consuming and so require a much greater investment on the part of the researcher and the participants. A third approach, known as **cross-sequential studies**, combines elements of both cross-sectional and longitudinal studies. Rather than measuring differences between people in different age groups or following the same people over a long period of time, researchers adopting this approach choose a smaller period of time during which they follow people in different age groups. For example, they might measure changes over a ten year period among participants who at the start of the study fall into the following age groups: 20 years old, 30 years old, 40 years old, 50 years old, and 60 years old. This design is advantageous because the researcher reaps the immediate benefits of being able to compare the age groups after the first assessment. Further, by following the different age groups over time they can subsequently determine whether the original differences they found across the age groups are due to true age effects or cohort effects.

The types of research we have discussed so far are all quantitative, referring to the fact that the data consist of numbers that are analyzed using statistical techniques. But as you will learn in this chapter, many observational research studies are more qualitative in nature. In qualitative research, the data are usually nonnumerical and therefore cannot be analyzed using statistical techniques. Rosenhan's observational study of the experience of people in psychiatric wards was primarily qualitative. The data were the notes taken by the "pseudopatients"—the people pretending to have heard voices—along with their hospital records. Rosenhan's analysis consists mainly of a written description of the experiences of the pseudopatients, supported by several concrete examples. To illustrate the hospital staff's tendency to "depersonalize" their patients, he noted, "Upon being admitted, I and other pseudopatients took the initial physical examinations in a semi-public room, where staff members went about their own business as if we were not there" (Rosenhan, 1973, p. 256)[2]. Qualitative data has a separate set of analysis tools depending on the research question. For example, thematic analysis would focus on themes that emerge in the data or conversation analysis would focus on the way the words were said in an interview or focus group.

Internal Validity Revisited

Recall that internal validity is the extent to which the design of a study supports the conclusion that changes in the independent variable caused any observed differences in the dependent variable. Figure 6.1 shows how experimental,

2. Rosenhan, D. L. (1973). On being sane in insane places. *Science*, *179*, 250–258.

quasi-experimental, and non-experimental (correlational) research vary in terms of internal validity. Experimental research tends to be highest in internal validity because the use of manipulation (of the independent variable) and control (of extraneous variables) help to rule out alternative explanations for the observed relationships. If the average score on the dependent variable in an experiment differs across conditions, it is quite likely that the independent variable is responsible for that difference. Non-experimental (correlational) research is lowest in internal validity because these designs fail to use manipulation or control. Quasi-experimental research (which will be described in more detail in a subsequent chapter) falls in the middle because it contains some, but not all, of the features of a true experiment. For instance, it may fail to use random assignment to assign participants to groups or fail to use counterbalancing to control for potential order effects. Imagine, for example, that a researcher finds two similar schools, starts an anti-bullying program in one, and then finds fewer bullying incidents in that "treatment school" than in the "control school." While a comparison is being made with a control condition, the inability to randomly assign children to schools could still mean that students in the treatment school differed from students in the control school in some other way that could explain the difference in bullying (e.g., there may be a selection effect).

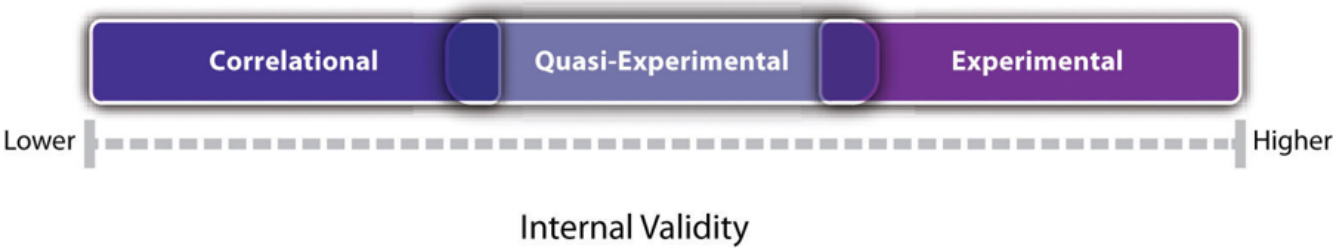


Figure 6.1 Internal Validity of Correlation, Quasi-Experimental, and Experimental Studies. Experiments are generally high in internal validity, quasi-experiments lower, and correlation (non-experimental) studies lower still.

Notice also in Figure 6.1 that there is some overlap in the internal validity of experiments, quasi-experiments, and correlational (non-experimental) studies. For example, a poorly designed experiment that includes many confounding variables can be lower in internal validity than a well-designed quasi-experiment with no obvious confounding variables. Internal validity is also only one of several validities that one might consider, as noted in Chapter 5.

29. Correlational Research

Learning Objectives

1. Define correlational research and give several examples.
2. Explain why a researcher might choose to conduct correlational research rather than experimental research or another type of non-experimental research.
3. Interpret the strength and direction of different correlation coefficients.
4. Explain why correlation does not imply causation.

What Is Correlational Research?

Correlational research is a type of non-experimental research in which the researcher measures two variables (binary or continuous) and assesses the statistical relationship (i.e., the correlation) between them with little or no effort to control extraneous variables. There are many reasons that researchers interested in statistical relationships between variables would choose to conduct a correlational study rather than an experiment. The first is that they do not believe that the statistical relationship is a causal one or are not interested in causal relationships. Recall two goals of science are to describe and to predict and the correlational research strategy allows researchers to achieve both of these goals. Specifically, this strategy can be used to describe the strength and direction of the relationship between two variables and if there is a relationship between the variables then the researchers can use scores on one variable to predict scores on the other (using a statistical technique called regression, which is discussed further in the section on Complex Correlation in this chapter).

Another reason that researchers would choose to use a correlational study rather than an experiment is that the statistical relationship of interest is thought to be causal, but the researcher *cannot* manipulate the independent variable because it is impossible, impractical, or unethical. For example, while a researcher might be interested in the relationship between the frequency people use cannabis and their memory abilities they cannot ethically manipulate the frequency that people use cannabis. As such, they must rely on the correlational research strategy; they must simply measure the frequency that people use cannabis and measure their memory abilities using a standardized test of memory and then determine whether the frequency people use cannabis is statistically related to memory test performance.

Correlation is also used to establish the reliability and validity of measurements. For example, a researcher might evaluate the validity of a brief extraversion test by administering it to a large group of participants along with a longer extraversion test that has already been shown to be valid. This researcher might then check to see whether participants' scores on the brief test are strongly correlated with their scores on the longer one. Neither test score is thought to cause the other, so there is no independent variable to manipulate. In fact, the terms *independent variable* and *dependent variable* do not apply to this kind of research.

Another strength of correlational research is that it is often higher in external validity than experimental research. Recall there is typically a trade-off between internal validity and external validity. As greater controls are added to experiments, internal validity is increased but often at the expense of external validity as artificial conditions are introduced that do not exist in reality. In contrast, correlational studies typically have low internal validity because

nothing is manipulated or controlled but they often have high external validity. Since nothing is manipulated or controlled by the experimenter the results are more likely to reflect relationships that exist in the real world.

Finally, extending upon this trade-off between internal and external validity, correlational research can help to provide converging evidence for a theory. If a theory is supported by a true experiment that is high in internal validity as well as by a correlational study that is high in external validity then the researchers can have more confidence in the validity of their theory. As a concrete example, correlational studies establishing that there is a relationship between watching violent television and aggressive behavior have been complemented by experimental studies confirming that the relationship is a causal one (Bushman & Huesmann, 2001)[1].

Does Correlational Research Always Involve Quantitative Variables?

A common misconception among beginning researchers is that correlational research must involve two quantitative variables, such as scores on two extraversion tests or the number of daily hassles and number of symptoms people have experienced. However, the defining feature of correlational research is that the two variables are measured—neither one is manipulated—and this is true regardless of whether the variables are quantitative or categorical. Imagine, for example, that a researcher administers the Rosenberg Self-Esteem Scale to 50 American college students and 50 Japanese college students. Although this "feels" like a between-subjects experiment, it is a correlational study because the researcher did not manipulate the students' nationalities. The same is true of the study by Cacioppo and Petty comparing college faculty and factory workers in terms of their need for cognition. It is a correlational study because the researchers did not manipulate the participants' occupations.

Figure 6.2 shows data from a hypothetical study on the relationship between whether people make a daily list of things to do (a "to-do list") and stress. Notice that it is unclear whether this is an experiment or a correlational study because it is unclear whether the independent variable was manipulated. If the researcher randomly assigned some participants to make daily to-do lists and others not to, then it is an experiment. If the researcher simply asked participants whether they made daily to-do lists, then it is a correlational study. The distinction is important because if the study was an experiment, then it could be concluded that making the daily to-do lists reduced participants' stress. But if it was a correlational study, it could only be concluded that these variables are statistically related. Perhaps being stressed has a negative effect on people's ability to plan ahead (the directionality problem). Or perhaps people who are more conscientious are more likely to make to-do lists and less likely to be stressed (the third-variable problem). The crucial point is that what defines a study as experimental or correlational is not the variables being studied, nor whether the variables are quantitative or categorical, nor the type of graph or statistics used to analyze the data. What defines a study is *how* the study is conducted.

1. Bushman, B. J., & Huesmann, L. R. (2001). Effects of televised violence on aggression. In D. Singer & J. Singer (Eds.), *Handbook of children and the media* (pp. 223–254). Thousand Oaks, CA: Sage.

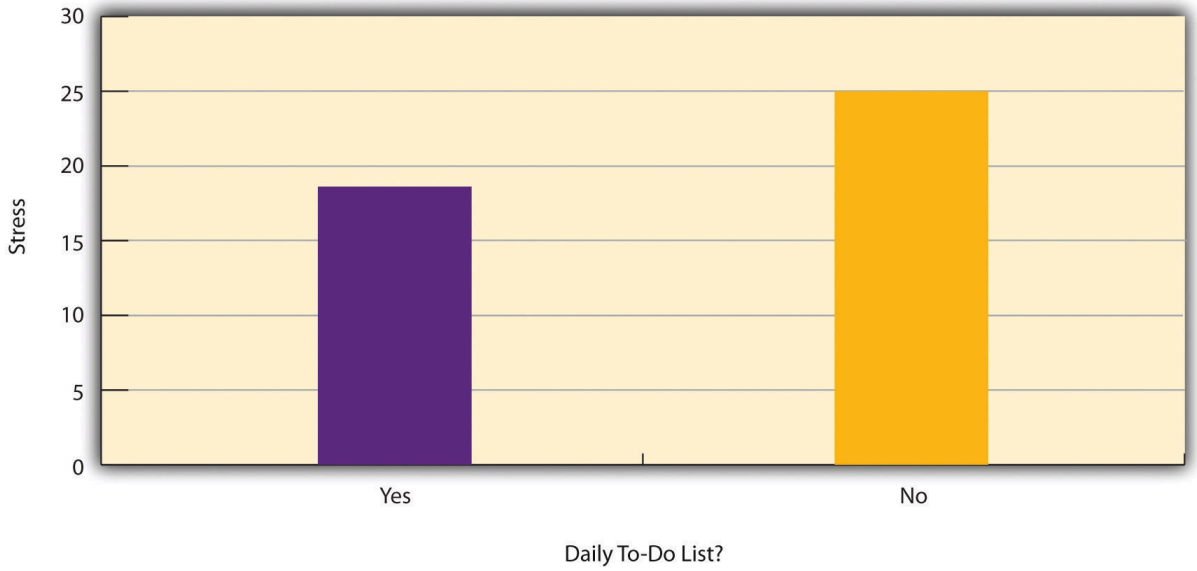


Figure 6.2 Results of a Hypothetical Study on Whether People Who Make Daily To-Do Lists Experience Less Stress Than People Who Do Not Make Such Lists

Data Collection in Correlational Research

Again, the defining feature of correlational research is that neither variable is manipulated. It does not matter how or where the variables are measured. A researcher could have participants come to a laboratory to complete a computerized backward digit span task and a computerized risky decision-making task and then assess the relationship between participants' scores on the two tasks. Or a researcher could go to a shopping mall to ask people about their attitudes toward the environment and their shopping habits and then assess the relationship between these two variables. Both of these studies would be correlational because no independent variable is manipulated.

Correlations Between Quantitative Variables

Correlations between quantitative variables are often presented using **scatterplots**. Figure 6.3 shows some hypothetical data on the relationship between the amount of stress people are under and the number of physical symptoms they have. Each point in the scatterplot represents one person's score on both variables. For example, the circled point in Figure 6.3 represents a person whose stress score was 10 and who had three physical symptoms. Taking all the points into account, one can see that people under more stress tend to have more physical symptoms. This is a good example of a **positive relationship**, in which higher scores on one variable tend to be associated with higher scores on the other. In other words, they move in the same direction, either both up or both down. A **negative relationship** is one in which higher scores on one variable tend to be associated with lower scores on the other. In other words, they move

in opposite directions. There is a negative relationship between stress and immune system functioning, for example, because higher stress is associated with lower immune system functioning.

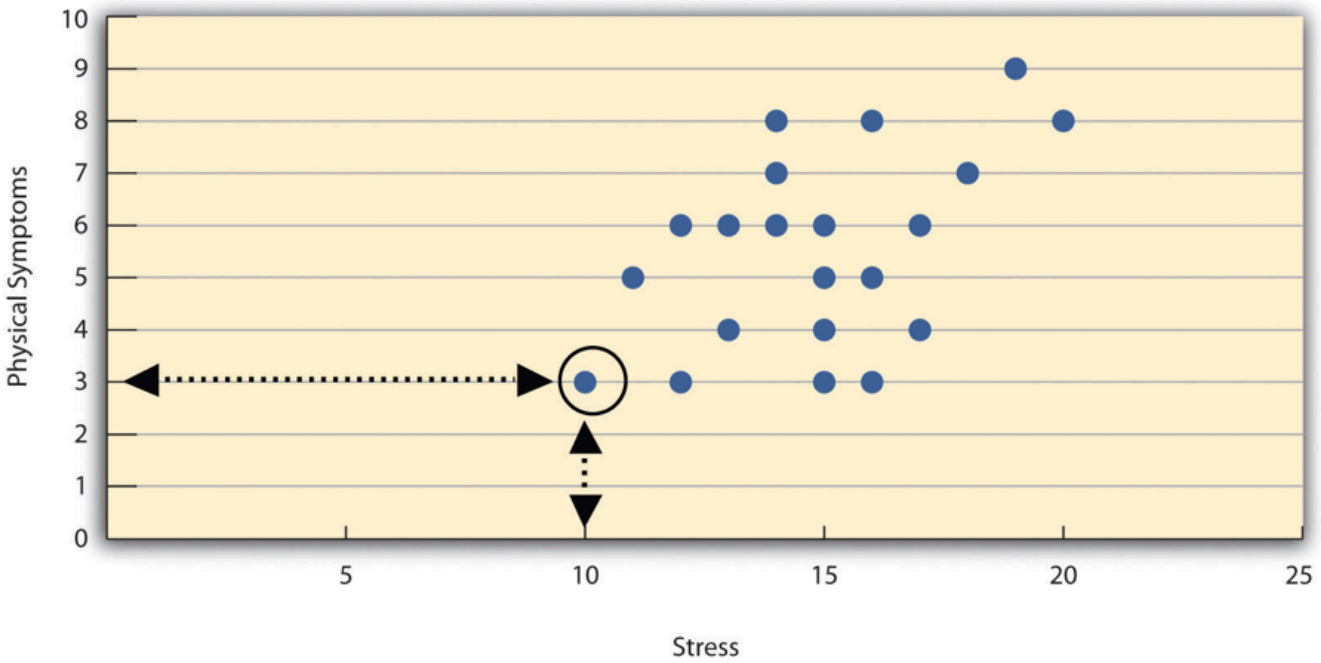


Figure 6.3 Scatterplot Showing a Hypothetical Positive Relationship Between Stress and Number of Physical Symptoms. The circled point represents a person whose stress score was 10 and who had three physical symptoms. Pearson's r for these data is +.51.

The strength of a correlation between quantitative variables is typically measured using a statistic called **Pearson's Correlation Coefficient (or Pearson's *r*)**. As Figure 6.4 shows, Pearson's r ranges from –1.00 (the strongest possible negative relationship) to +1.00 (the strongest possible positive relationship). A value of 0 means there is no relationship between the two variables. When Pearson's *r* is 0, the points on a scatterplot form a shapeless "cloud." As its value moves toward –1.00 or +1.00, the points come closer and closer to falling on a single straight line. Correlation coefficients near ±.10 are considered small, values near ± .30 are considered medium, and values near ±.50 are considered large. Notice that the sign of Pearson's *r* is unrelated to its strength. Pearson's *r* values of +.30 and –.30, for example, are equally strong; it is just that one represents a moderate positive relationship and the other a moderate negative relationship. With the exception of reliability coefficients, most correlations that we find in Psychology are small or moderate in size. The website http://rpsychologist.com/d3/correlation/, created by Kristoffer Magnusson, provides an excellent interactive visualization of correlations that permits you to adjust the strength and direction of a correlation while witnessing the corresponding changes to the scatterplot.

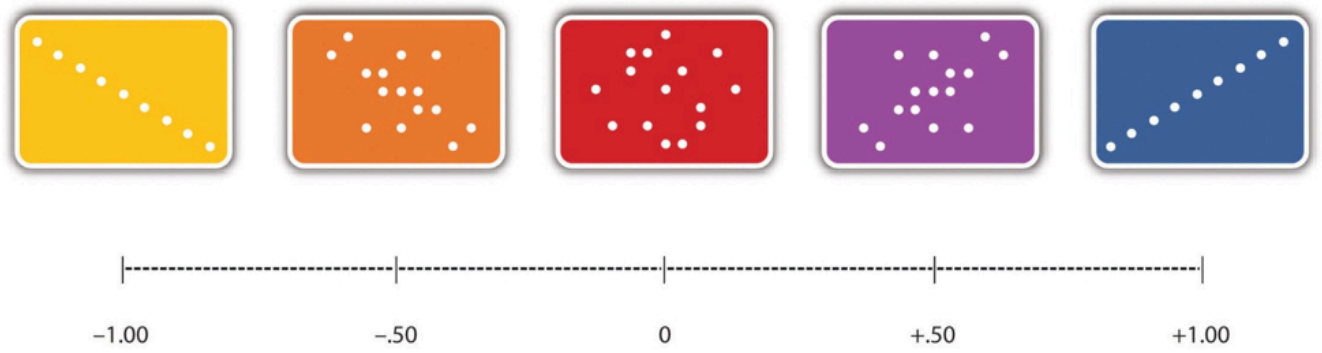


Figure 6.4 Range of Pearson's r, From –1.00 (Strongest Possible Negative Relationship), Through 0 (No Relationship), to +1.00 (Strongest Possible Positive Relationship)

There are two common situations in which the value of Pearson's *r* can be misleading. Pearson's *r* is a good measure only for linear relationships, in which the points are best approximated by a straight line. It is not a good measure for nonlinear relationships, in which the points are better approximated by a curved line. Figure 6.5, for example, shows a hypothetical relationship between the amount of sleep people get per night and their level of depression. In this example, the line that best approximates the points is a curve—a kind of upside-down "U"—because people who get about eight hours of sleep tend to be the least depressed. Those who get too little sleep and those who get too much sleep tend to be more depressed. Even though Figure 6.5 shows a fairly strong relationship between depression and sleep, Pearson's *r* would be close to zero because the points in the scatterplot are not well fit by a single straight line. This means that it is important to make a scatterplot and confirm that a relationship is approximately linear before using Pearson's *r*. Nonlinear relationships are fairly common in psychology, but measuring their strength is beyond the scope of this book.

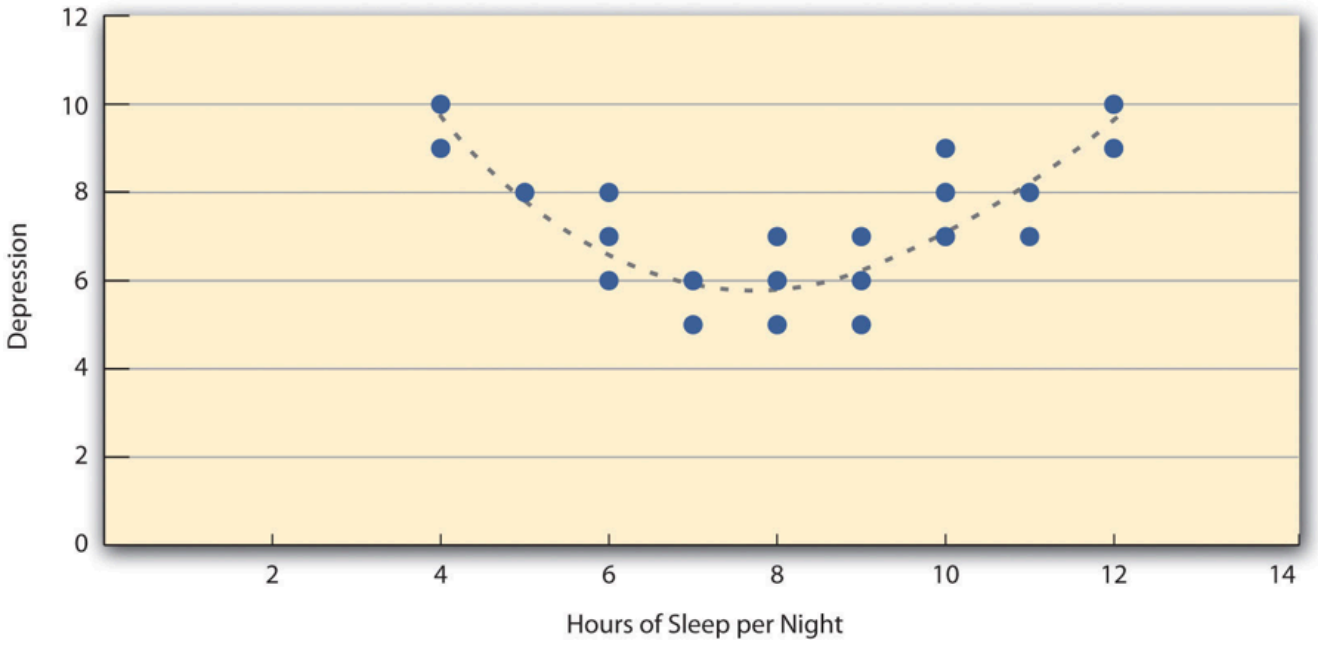


Figure 6.5 Hypothetical Nonlinear Relationship Between Sleep and Depression

The other common situations in which the value of Pearson's *r* can be misleading is when one or both of the variables

have a limited range in the sample relative to the population. This problem is referred to as **restriction of range**. Assume, for example, that there is a strong negative correlation between people's age and their enjoyment of hip hop music as shown by the scatterplot in Figure 6.6. Pearson's *r* here is –.77. However, if we were to collect data only from 18- to 24-year-olds—represented by the shaded area of Figure 6.6—then the relationship would seem to be quite weak. In fact, Pearson's *r* for this restricted range of ages is 0. It is a good idea, therefore, to design studies to avoid restriction of range. For example, if age is one of your primary variables, then you can plan to collect data from people of a wide range of ages. Because restriction of range is not always anticipated or easily avoidable, however, it is good practice to examine your data for possible restriction of range and to interpret Pearson's *r* in light of it. (There are also statistical methods to correct Pearson's *r* for restriction of range, but they are beyond the scope of this book).

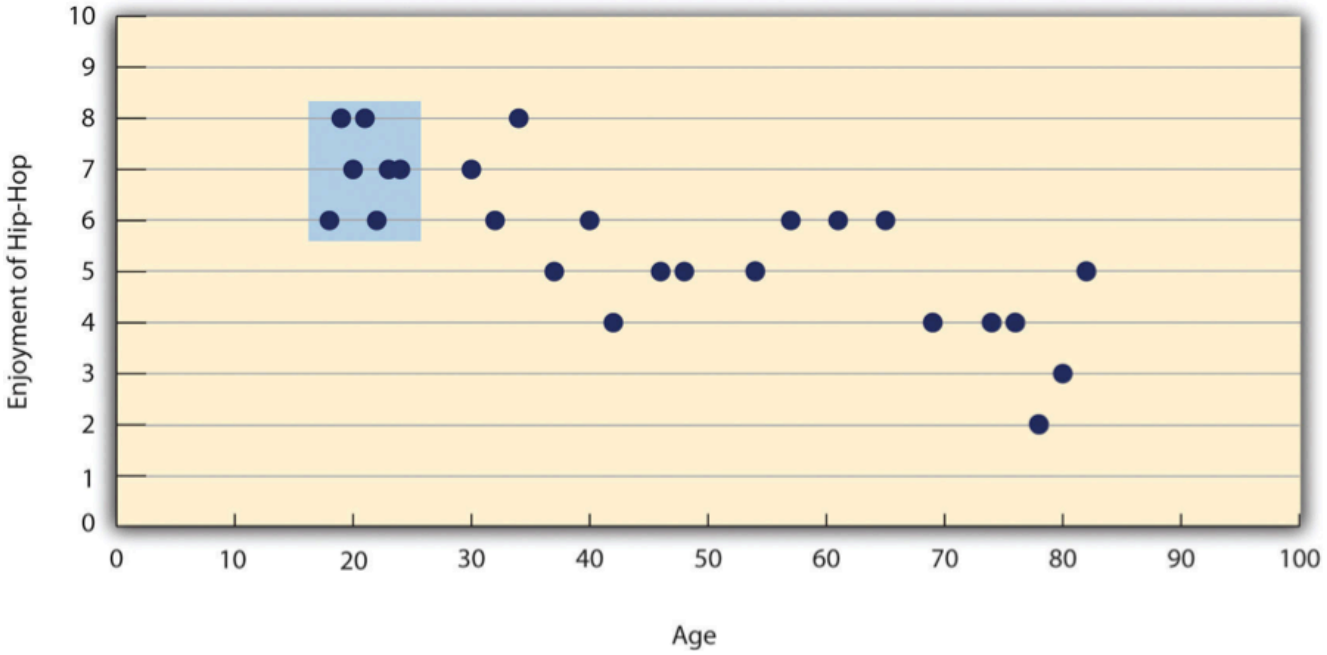


Figure 6.6 Hypothetical Data Showing How a Strong Overall Correlation Can Appear to Be Weak When One Variable Has a Restricted Range. The overall correlation here is –.77, but the correlation for the 18- to 24-year-olds (in the blue box) is 0.

Correlation Does Not Imply Causation

You have probably heard repeatedly that "Correlation does not imply causation." An amusing example of this comes from a 2012 study that showed a positive correlation (Pearson's $r = 0.79$) between the per capita chocolate consumption of a nation and the number of Nobel prizes awarded to citizens of that nation[2]. It seems clear, however, that this does not mean that eating chocolate causes people to win Nobel prizes, and it would not make sense to try to increase the number of Nobel prizes won by recommending that parents feed their children more chocolate.

There are two reasons that correlation does not imply causation. The first is called the **directionality problem**. Two variables, X and Y, can be statistically related because X causes Y or because Y causes X. Consider, for example, a study showing that whether or not people exercise is statistically related to how happy they are—such that people who exercise are happier on average than people who do not. This statistical relationship is consistent with the idea that

2. Messerli, F. H. (2012). Chocolate consumption, cognitive function, and Nobel laureates. *New England Journal of Medicine*, 367, 1562-1564.

exercising causes happiness, but it is also consistent with the idea that happiness causes exercise. Perhaps being happy gives people more energy or leads them to seek opportunities to socialize with others by going to the gym. The second reason that correlation does not imply causation is called the **third-variable problem**. Two variables, X and Y, can be statistically related not because X causes Y, or because Y causes X, but because some third variable, Z, causes both X and Y. For example, the fact that nations that have won more Nobel prizes tend to have higher chocolate consumption probably reflects geography in that European countries tend to have higher rates of per capita chocolate consumption *and* invest more in education and technology (once again, per capita) than many other countries in the world. Similarly, the statistical relationship between exercise and happiness could mean that some third variable, such as physical health, causes both of the others. Being physically healthy could cause people to exercise and cause them to be happier. Correlations that are a result of a third-variable are often referred to as **spurious correlations.**

Some excellent and amusing examples of spurious correlations can be found at http://www.tylervigen.com (Figure 6.7 provides one such example).

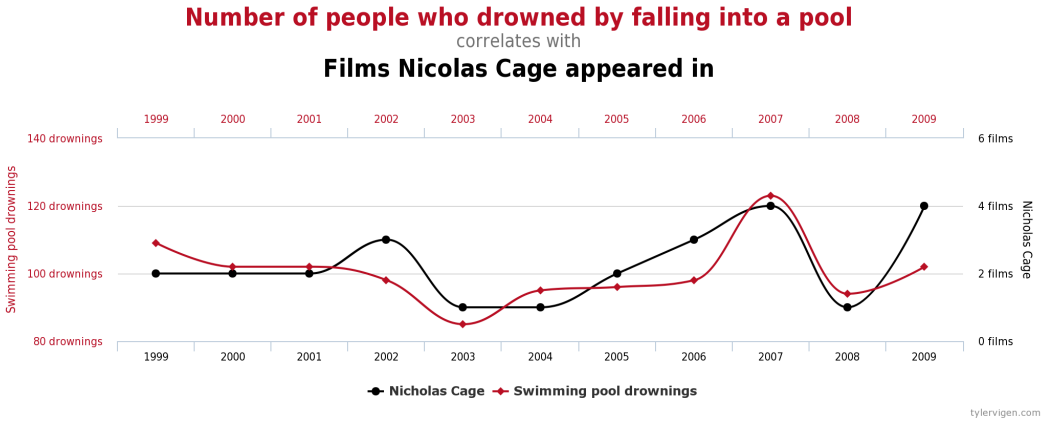


Figure 6.7 Example of a Spurious Correlation. Source: http://tylervigen.com/spurious-correlations (CC-BY 4.0)

"Lots of Candy Could Lead to Violence"

Although researchers in psychology know that correlation does not imply causation, many journalists do not. One website about correlation and causation, http://jonathan.mueller.faculty.noctrl.edu/100/correlation_or_causation.htm, links to dozens of media reports about real biomedical and psychological research. Many of the headlines suggest that a causal relationship has been demonstrated when a careful reading of the articles shows that it has not because of the directionality and third-variable problems.

One such article is about a study showing that children who ate candy every day were more likely than other children to be arrested for a violent offense later in life. But could candy really "lead to" violence, as the headline suggests? What alternative explanations can you think of for this statistical relationship? How could the headline be rewritten so that it is not misleading?

As you have learned by reading this book, there are various ways that researchers address the directionality and third-variable problems. The most effective is to conduct an experiment. For example, instead of simply measuring how much people exercise, a researcher could bring people into a laboratory and randomly assign half of them to run on a treadmill for 15 minutes and the rest to sit on a couch for 15 minutes. Although this seems like a minor change to the research design, it is extremely important. Now if the exercisers end up in

more positive moods than those who did not exercise, it cannot be because their moods affected how much they exercised (because it was the *researcher* who used random assignment to determine how much they exercised). Likewise, it cannot be because some third variable (e.g., physical health) affected both how much they exercised and what mood they were in. Thus experiments eliminate the directionality and third-variable problems and allow researchers to draw firm conclusions about causal relationships.

30. Complex Correlation

Learning Objectives

1. Explain some reasons that researchers use complex correlational designs.
2. Create and interpret a correlation matrix.
3. Describe how researchers can use partial correlation and multiple regression to statistically control for third variables.

As we have already seen, researchers conduct correlational studies rather than experiments when they are interested in noncausal relationships or when they are interested in causal relationships but the independent variable cannot be manipulated for practical or ethical reasons. In this section, we look at some approaches to complex correlational research that involve measuring several variables and assessing the relationships among them.

Assessing Relationships Among Multiple Variables

Most complex correlational research involves measuring several variables—either binary or continuous—and then assessing the statistical relationships among them. For example, researchers Nathan Radcliffe and William Klein studied a sample of middle-aged adults to see how their level of optimism (measured by using a short questionnaire called the Life Orientation Test) relates to several other variables related to having a heart attack (Radcliffe & Klein, 2002)[1]. These included their health, their knowledge of heart attack risk factors, and their beliefs about their own risk of having a heart attack. They found that more optimistic participants were healthier (e.g., they exercised more and had lower blood pressure), knew about heart attack risk factors, and correctly believed their own risk to be lower than that of their peers.

In another example, Ernest Jouriles and his colleagues measured adolescents' experiences of physical and psychological relationship aggression and their psychological distress. Because measures of physical aggression (such as the Conflict in Adolescent Dating Relationships Inventory and the Relationship Violence Interview) often tend to result in highly skewed distributions, the researchers transformed their measures of physical aggression into a dichotomous (i.e., binary) measure (0 = did not occur, 1 = did occur). They did the same with their measures of psychological aggression and then measured the correlations among these variables, finding that adolescents who experienced physical aggression were moderately likely to also have experienced psychological aggression and that experiencing psychological aggression was related to symptoms of psychological distress. (Jouriles, Garrido, Rosenfield, & McDonald, 2009)[2]

1. Radcliffe, N. M., & Klein, W. M. P. (2002). Dispositional, unrealistic, and comparative optimism: Differential relations with knowledge and processing of risk information and beliefs about personal risk. *Personality and Social Psychology Bulletin*, *28*, 836–846.
2. Jouriles, E. N., Garrido, E., Rosenfield, D., & McDonald, R. (2009). Experiences of psychological and

This approach is often used to assess the validity of new psychological measures. For example, when John Cacioppo and Richard Petty created their Need for Cognition Scale—a measure of the extent to which people like to think and value thinking—they used it to measure the need for cognition for a large sample of college students, along with three other variables: intelligence, socially desirable responding (the tendency to give what one thinks is the "appropriate" response), and dogmatism (Caccioppo & Petty, 1982)[3]. The results of this study are summarized in Table 6.1, which is a **correlation matrix** showing the correlation (Pearson's *r*) between every possible pair of variables in the study. For example, the correlation between the need for cognition and intelligence was +.39, the correlation between intelligence and socially desirable responding was +.02, and so on. (Only half the matrix is filled in because the other half would contain exactly the same information. Also, because the correlation between a variable and itself is always +1.00, these values are replaced with dashes throughout the matrix.) In this case, the overall pattern of correlations was consistent with the researchers' ideas about how scores on the need for cognition should be related to these other constructs.

Table 6.1 Correlation Matrix Showing Correlations Among the Need for Cognition and Three Other Variables Based on Research by Cacioppo and Petty (1982)

	Need for cognition	Intelligence	Social desirability	Dogmatism
Need for cognition	–			
Intelligence	+.39	–		
Social desirability	+.08	+.02	–	
Dogmatism	–.27	–.23	+.03	–

Factor Analysis

When researchers study relationships among a large number of conceptually similar variables, they often use a complex statistical technique called **factor analysis**. In essence, factor analysis organizes the variables into a smaller number of clusters, such that they are strongly correlated within each cluster but weakly correlated between clusters. Each cluster is then interpreted as multiple measures of the same underlying construct. These underlying constructs are also called "factors." For example, when people perform a wide variety of mental tasks, factor analysis typically organizes them into two main factors—one that researchers interpret as mathematical intelligence (arithmetic, quantitative estimation, spatial reasoning, and so on) and another that they interpret as verbal intelligence (grammar, reading comprehension, vocabulary, and so on). The Big Five personality factors have been identified through factor analyses of people's scores on a large number of more specific traits. For example, measures of warmth, gregariousness, activity level, and positive emotions tend to be highly correlated with each other and are interpreted as representing the construct of extraversion. As a final example, researchers Peter Rentfrow and Samuel Gosling asked more than 1,700 university students to rate how much they liked 14 different popular genres of music (Rentfrow & Gosling, 2008)[4]. They then submitted these 14 variables to a factor analysis, which identified four distinct factors. The researchers called them *Reflective and Complex* (blues, jazz, classical, and folk), *Intense and Rebellious* (rock, alternative, and heavy

physical aggression in adolescent romantic relationships: Links to psychological distress. *Child Abuse & Neglect*, 33(7), 451–460.

3. Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116–131.
4. Rentfrow, P. J., & Gosling, S. D. (2008). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84, 1236–1256.

metal), *Upbeat and Conventional* (country, soundtrack, religious, pop), and *Energetic and Rhythmic* (rap/hip-hop, soul/funk, and electronica); see Table 6.2.

Table 6.2 Factor Loadings of the 14 Music Genres on Four Varimax-Rotated Principal Components. Based on Research by Rentfrow and Gosling (2003)

	Music-preference dimension			
Genre	***Reflective and Complex***	***Intense and Rebellious***	***Upbeat and Conventional***	***Energetic and Rhythmic***
Blues	**.85**	.01	-.09	.12
Jazz	**.83**	.04	.07	.15
Classical	**.66**	.14	.02	-.13
Folk	**.64**	.09	.15	-.16
Rock	.17	**.85**	-.04	-.07
Alternative	.02	**.80**	.13	.04
Heavy metal	.07	**.75**	-.11	.04
Country	-.06	.05	**.72**	-.03
Sound tracks	.01	.04	**.70**	.17
Religious	.23	-.21	**.64**	-.01
Pop	-.20	.06	**.59**	.45
Rap/hip-hop	-.19	-.12	.17	**.79**
Soul/funk	.39	-.11	.11	**.69**
Electronica/dance	-.02	.15	-.01	**.60**

Note. N = 1,704. All factor loadings .40 or larger are in italics; the highest factor loadings for each dimension are listed in boldface type.

Two additional points about factor analysis are worth making here. One is that factors are not categories. Factor analysis does not tell us that people are *either* extraverted *or* conscientious or that they like *either* "reflective and complex" music *or* "intense and rebellious" music. Instead, factors are constructs that operate independently of each other. So people who are high in extraversion might be high or low in conscientiousness, and people who like reflective and complex music might or might not also like intense and rebellious music. The second point is that factor analysis reveals only the underlying structure of the variables. It is up to researchers to interpret and label the factors and to explain the origin of that particular factor structure. For example, one reason that extraversion and the other Big Five operate as separate factors is that they appear to be controlled by different genes (Plomin, DeFries, McClean, & McGuffin, 2008)[5].

Exploring Causal Relationships

Another important use of complex correlational research is to explore possible causal relationships among variables. This might seem surprising given the oft-quoted saying that "correlation does not imply causation." It is true that correlational research cannot unambiguously establish that one variable causes another. Complex correlational

5. Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2008). *Behavioral genetics* (5th ed.). New York, NY: Worth.

research, however, can often be used to rule out other plausible interpretations. The primary way of doing this is through the **statistical control** of potential third variables. Instead of controlling these variables through random assignment or by holding them constant as in an experiment, the researcher instead measures them and includes them in the statistical analysis called **partial correlation**. Using this technique, researchers can examine the relationship between two variables, while statistically controlling for one or more potential third variables.

For example, assume a researcher was interested in the relationship between watching violent television shows and aggressive behavior but she was concerned that socioeconomic status (SES) might represent a third variable that is driving this relationship. In this case, she could conduct a study in which she measures the amount of violent television that participants watch in their everyday life, the number of acts of aggression that they have engaged in, and their SES. She could first examine the correlation between violent television viewing and aggression. Let's say she found a correlation of +.35, which would be considered a moderate sized positive correlation. Next, she could use partial correlation to reexamine this relationship after statistically controlling for SES. This technique would allow her to examine the relationship between the part of violent television viewing that is independent of SES and the part of aggressive behavior that is independent of SES. If she found that the partial correlation between violent television viewing and aggression while controlling for SES was +.34, that would suggest that the relationship between violent television viewing and aggression is largely independent of SES (i.e., SES is not a third variable driving this relationship). On the other hand, if she found that after statistically controlling for SES the correlation between violent television viewing and aggression dropped to +.03, then that would suggest that SES is indeed a third variable that is driving the relationship. If, however, she found that statistically controlling for SES reduced the magnitude of the correlation from +.35 to +.20, then this would suggest that SES accounts for some, but not all, of the relationship between television violence and aggression. It is important to note that while partial correlation provides an important tool for researchers to statistically control for third variables, researchers using this technique are still limited in their ability to arrive at causal conclusions because this technique does not take care of the directionality problem and there may be other third variables driving the relationship that the researcher did not consider and statistically control.

Regression

Once a relationship between two variables has been established, researchers can use that information to make predictions about the value of one variable given the value of another variable. For, instance, once we have established that there is a correlation between IQ and GPA we can use people's IQ scores to predict their GPA. Thus, while correlation coefficients can be used to describe the strength and direction of relationships between variables, **regression** is a statistical technique that allows researchers to predict one variable given another. Regression can also be used to describe more complex relationships between more than two variables. Typically the variable that is used to make the prediction is referred to as the **predictor variable** and the variable that is being predicted is called the **outcome variable or criterion variable.** This regression equation has the following general form:

$Y = b_1X_1$

Y in this formula represents the person's predicted score on the outcome variable, b_1 represents the slope of the line depicting the relationship between two variables (or the regression weight), and X_1 represents the person's score on the predictor variable. You can see that to predict a person's score on the outcome variable (Y), one simply needs to multiply their score on the predictor variable (X) by the regression weight (b_1)

While **simple regression** involves using one variable to predict another, **multiple regression** involves measuring several variables (X1, X2, X3,…Xi), and using them to predict some outcome variable (Y). Multiple regression can also be used to simply describe the relationship between a single outcome variable (Y) and a set of predictor variables (X1, X2, X3,…Xi). The result of a multiple regression analysis is an equation that expresses the outcome variable as an additive combination of the predictor variables. This regression equation has the following general form:

$Y = b_1X_1 + b_2X_2 + b_3X_3 + \ldots + b_iXi$

The regression weights (b_1, b_2, and so on) indicate how large a contribution a predictor variable makes, on average, to the prediction of the outcome variable. Specifically, they indicate how much the outcome variable changes for each one-unit change in the predictor variable.

The advantage of multiple regression is that it can show whether a predictor variable makes a contribution to an outcome variable *over and above* the contributions made by other predictor variables (i.e., it can be used to show whether a predictor variable is related to an outcome variable after statistically controlling for other predictor variables). As a hypothetical example, imagine that a researcher wants to know how income and health relate to happiness. This is tricky because income and health are themselves related to each other. Thus if people with greater incomes tend to be happier, then perhaps this is only because they tend to be healthier. Likewise, if people who are healthier tend to be happier, perhaps this is only because they tend to make more money. But a multiple regression analysis including both income and health as predictor variables would show whether each one makes a contribution to the prediction of happiness when the other is taken into account (when it is statistically controlled). In other words, multiple regression would allow the researcher to examine whether that part of income that is unrelated to health predicts or relates to happiness as well as whether that part of health that is unrelated to income predicts or relates to happiness. Research like this, by the way, has shown both income and health make extremely small contributions to happiness except in the case of severe poverty or illness (Diener, 2000[6]).

The examples discussed in this section only scratch the surface of how researchers use complex correlational research to explore possible causal relationships among variables. It is important to keep in mind, however, that purely correlational approaches cannot unambiguously establish that one variable causes another. The best they can do is show patterns of relationships that are consistent with some causal interpretations and inconsistent with others.

6. Diener, E. (2000). Subjective well-being: The science of happiness, and a proposal for a national index. *American Psychologist*, 55, 34–43.

31. Qualitative Research

Learning Objectives

1. List several ways in which qualitative research differs from quantitative research in psychology.
2. Describe the strengths and weaknesses of qualitative research in psychology compared with quantitative research.
3. Give examples of qualitative research in psychology.

What Is Qualitative Research?

This textbook is primarily about **quantitative research,** in part because most studies conducted in psychology are quantitative in nature. Quantitative researchers typically start with a focused research question or hypothesis, collect a small amount of numerical data from a large number of individuals, describe the resulting data using statistical techniques, and draw general conclusions about some large population. Although this method is by far the most common approach to conducting empirical research in psychology, there is an important alternative called **qualitative research**. Qualitative research originated in the disciplines of anthropology and sociology but is now used to study psychological topics as well. Qualitative researchers generally begin with a less focused research question, collect large amounts of relatively "unfiltered" data from a relatively small number of individuals, and describe their data using nonstatistical techniques, such as grounded theory, thematic analysis, critical discourse analysis, or interpretative phenomenological analysis. They are usually less concerned with drawing general conclusions about human behavior than with understanding in detail the *experience* of their research participants.

Consider, for example, a study by researcher Per Lindqvist and his colleagues, who wanted to learn how the families of teenage suicide victims cope with their loss (Lindqvist, Johansson, & Karlsson, 2008)[1]. They did not have a specific research question or hypothesis, such as, What percentage of family members join suicide support groups? Instead, they wanted to understand the variety of reactions that families had, with a focus on what it is like from *their* perspectives. To address this question, they interviewed the families of 10 teenage suicide victims in their homes in rural Sweden. The interviews were relatively unstructured, beginning with a general request for the families to talk about the victim and ending with an invitation to talk about anything else that they wanted to tell the interviewer. One of the most important themes that emerged from these interviews was that even as life returned to "normal," the families continued to struggle with the question of why their loved one committed suicide. This struggle appeared to be especially difficult for families in which the suicide was most unexpected.

1. Lindqvist, P., Johansson, L., & Karlsson, U. (2008). In the aftermath of teenage suicide: A qualitative study of the psychosocial consequences for the surviving family members. *BMC Psychiatry*, 8, 26. Retrieved from http://www.biomedcentral.com/1471-244X/8/26

The Purpose of Qualitative Research

Again, this textbook is primarily about quantitative research in psychology. The strength of quantitative research is its ability to provide precise answers to specific research questions and to draw general conclusions about human behavior. This method is how we know that people have a strong tendency to obey authority figures, for example, and that female undergraduate students are not substantially more talkative than male undergraduate students. But while quantitative research is good at providing precise answers to specific research questions, it is not nearly as good at *generating* novel and interesting research questions. Likewise, while quantitative research is good at drawing general conclusions about human behavior, it is not nearly as good at providing detailed descriptions of the behavior of particular groups in particular situations. And quantitative research is not very good at communicating what it is actually like to be a member of a particular group in a particular situation.

But the relative weaknesses of quantitative research are the relative strengths of qualitative research. Qualitative research can help researchers to generate new and interesting research questions and hypotheses. The research of Lindqvist and colleagues, for example, suggests that there may be a general relationship between how unexpected a suicide is and how consumed the family is with trying to understand why the teen committed suicide. This relationship can now be explored using quantitative research. But it is unclear whether this question would have arisen at all without the researchers sitting down with the families and listening to what they themselves wanted to say about their experience. Qualitative research can also provide rich and detailed descriptions of human behavior in the real-world contexts in which it occurs. Among qualitative researchers, this depth is often referred to as "thick description" (Geertz, 1973)[2]. Similarly, qualitative research can convey a sense of what it is actually like to be a member of a particular group or in a particular situation—what qualitative researchers often refer to as the "lived experience" of the research participants. Lindqvist and colleagues, for example, describe how all the families spontaneously offered to show the interviewer the victim's bedroom or the place where the suicide occurred—revealing the importance of these physical locations to the families. It seems unlikely that a quantitative study would have discovered this detail.

Table 6.3 Some contrasts between qualitative and quantitative research

Qualitative	**Quantitative**
1. In-depth information about relatively few people	1. Less depth information with larger samples
2. Conclusions are based on interpretations drawn by the investigator	2. Conclusions are based on statistical analyses
3. Global and exploratory	3. Specific and focused

Data Collection and Analysis in Qualitative Research

Data collection approaches in qualitative research are quite varied and can involve naturalistic observation, participant observation, archival data, artwork, and many other things. But one of the most common approaches, especially for psychological research, is to conduct **interviews**. Interviews in qualitative research can be unstructured—consisting of a small number of general questions or prompts that allow participants to talk about what is of interest to them—or structured, where there is a strict script that the interviewer does not deviate from. Most interviews are in between the two and are called semi-structured interviews, where the researcher has a few consistent questions and can follow up by asking more detailed questions about the topics that come up. Such interviews can be lengthy and detailed, but they are usually conducted with a relatively small sample. The unstructured interview was the approach used by Lindqvist

2. Geertz, C. (1973). *The interpretation of cultures*. New York, NY: Basic Books.

and colleagues in their research on the families of suicide victims because the researchers were aware that how much was disclosed about such a sensitive topic should be led by the families, not by the researchers.

Another approach used in qualitative research involves small groups of people who participate together in interviews focused on a particular topic or issue, known as **focus groups**. The interaction among participants in a focus group can sometimes bring out more information than can be learned in a one-on-one interview. The use of focus groups has become a standard technique in business and industry among those who want to understand consumer tastes and preferences. The content of all focus group interviews is usually recorded and transcribed to facilitate later analyses. However, we know from social psychology that group dynamics are often at play in any group, including focus groups, and it is useful to be aware of those possibilities. For example, the desire to be liked by others can lead participants to provide inaccurate answers that they believe will be perceived favorably by the other participants. The same may be said for personality characteristics. For example, highly extraverted participants can sometimes dominate discussions within focus groups.

Data Analysis in Qualitative Research

Although quantitative and qualitative research generally differ along several important dimensions (e.g., the specificity of the research question, the type of data collected), it is the method of data analysis that distinguishes them more clearly than anything else. To illustrate this idea, imagine a team of researchers that conducts a series of unstructured interviews with people recovering from alcohol use disorder to learn about the role of their religious faith in their recovery. Although this project sounds like qualitative research, imagine further that once they collect the data, they code the data in terms of how often each participant mentions God (or a "higher power"), and they then use descriptive and inferential statistics to find out whether those who mention God more often are more successful in abstaining from alcohol. Now it sounds like quantitative research. In other words, the quantitative-qualitative distinction depends more on what researchers do with the data they have collected than with why or how they collected the data.

But what does qualitative data analysis look like? Just as there are many ways to collect data in qualitative research, there are many ways to analyze data. Here we focus on one general approach called **grounded theory** (Glaser & Strauss, 1967)[3]. This approach was developed within the field of sociology in the 1960s and has gradually gained popularity in psychology. Remember that in quantitative research, it is typical for the researcher to start with a theory, derive a hypothesis from that theory, and then collect data to test that specific hypothesis. In qualitative research using grounded theory, researchers start with the data and develop a theory or an interpretation that is "grounded in" those data. They do this analysis in stages. First, they identify ideas that are repeated throughout the data. Then they organize these ideas into a smaller number of broader themes. Finally, they write a **theoretical narrative**—an interpretation of the data in terms of the themes that they have identified. This theoretical narrative focuses on the subjective experience of the participants and is usually supported by many direct quotations from the participants themselves.

As an example, consider a study by researchers Laura Abrams and Laura Curran, who used the grounded theory approach to study the experience of postpartum depression symptoms among low-income mothers (Abrams & Curran, 2009)[4]. Their data were the result of unstructured interviews with 19 participants. Table 6.4 shows the five broad themes

3. Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago, IL: Aldine.
4. Abrams, L. S., & Curran, L. (2009). "And you're telling me not to stress?" A grounded theory study of postpartum depression symptoms among low-income mothers. *Psychology of Women Quarterly, 33*, 351–362.

the researchers identified and the more specific repeating ideas that made up each of those themes. In their research report, they provide numerous quotations from their participants, such as this one from "Destiny:"

> Well, just recently my apartment was broken into and the fact that his Medicaid for some reason was cancelled so a lot of things was happening within the last two weeks all at one time. So that in itself I don't want to say almost drove me mad but it put me in a funk....Like I really was depressed. (p. 357)

Their theoretical narrative focused on the participants' experience of their symptoms, not as an abstract "affective disorder" but as closely tied to the daily struggle of raising children alone under often difficult circumstances.

Table 6.4 Themes and Repeating Ideas in a Study of Postpartum Depression Among Low-Income Mothers. Based on Research by Abrams and Curran (2009).

Theme	Repeating ideas
Ambivalence	"I wasn't prepared for this baby," "I didn't want to have any more children."
Caregiving overload	"Please stop crying," "I need a break," "I can't do this anymore."
Juggling	"No time to breathe," "Everyone depends on me," "Navigating the maze."
Mothering alone	"I really don't have any help," "My baby has no father."
Real-life worry	"I don't have any money," "Will my baby be OK?" "It's not safe here."

The Quantitative-Qualitative "Debate"

Given their differences, it may come as no surprise that quantitative and qualitative research in psychology and related fields do not coexist in complete harmony. Some quantitative researchers criticize qualitative methods on the grounds that they lack objectivity, are difficult to evaluate in terms of reliability and validity, and do not allow generalization to people or situations other than those actually studied. At the same time, some qualitative researchers criticize quantitative methods on the grounds that they overlook the richness of human behavior and experience and instead answer simple questions about easily quantifiable variables.

In general, however, qualitative researchers are well aware of the issues of objectivity, reliability, validity, and generalizability. In fact, they have developed a number of frameworks for addressing these issues (which are beyond the scope of our discussion). And in general, quantitative researchers are well aware of the issue of oversimplification. They do not believe that all human behavior and experience can be adequately described in terms of a small number of variables and the statistical relationships among them. Instead, they use simplification as a strategy for uncovering general principles of human behavior.

Many researchers from both the quantitative and qualitative camps now agree that the two approaches can and should be combined into what has come to be called **mixed-methods research** (Todd, Nerlich, McKeown, & Clarke, 2004)[5]. (In fact, the studies by Lindqvist and colleagues and by Abrams and Curran both combined quantitative and qualitative approaches.) One approach to combining quantitative and qualitative research is to use qualitative research for hypothesis generation and quantitative research for hypothesis testing. Again, while a qualitative study might suggest that families who experience an unexpected suicide have more difficulty resolving the question of why, a well-designed quantitative study could test a hypothesis by measuring these specific variables in a large sample. A second approach to combining quantitative and qualitative research is referred to as **triangulation**. The idea is to use both quantitative and qualitative methods simultaneously to study the same general questions and to compare the results.

5. Todd, Z., Nerlich, B., McKeown, S., & Clarke, D. D. (2004) *Mixing methods in psychology: The integration of qualitative and quantitative methods in theory and practice*. London, UK: Psychology Press.

If the results of the quantitative and qualitative methods converge on the same general conclusion, they reinforce and enrich each other. If the results diverge, then they suggest an interesting new question: Why do the results diverge and how can they be reconciled?

Using qualitative research can often help clarify quantitative results via triangulation. Trenor, Yu, Waight, Zerda, and Sha (2008)[6] investigated the experience of female engineering students at a university. In the first phase, female engineering students were asked to complete a survey, where they rated a number of their perceptions, including their sense of belonging. Their results were compared across the student ethnicities, and statistically, the various ethnic groups showed no differences in their ratings of their sense of belonging. One might look at that result and conclude that ethnicity does not have anything to do with one's sense of belonging. However, in the second phase, the authors also conducted interviews with the students, and in those interviews, many minority students reported how the diversity of cultures at the university enhanced their sense of belonging. Without the qualitative component, we might have drawn the wrong conclusion about the quantitative results.

This example shows how qualitative and quantitative research work together to help us understand human behavior. Some researchers have characterized qualitative research as best for identifying behaviors or the phenomenon whereas quantitative research is best for understanding meaning or identifying the mechanism. However, Bryman (2012)[7] argues for breaking down the divide between these arbitrarily different ways of investigating the same questions.

6. Trenor, J.M., Yu, S.L., Waight, C.L., Zerda. K.S & Sha T.-L. (2008). The relations of ethnicity to female engineering students' educational experiences and college and career plans in an ethnically diverse learning environment. *Journal of Engineering Education*, 97(4), 449-465.
7. Bryman, A. (2012). *Social Research Methods*, 4th ed. Oxford: OUP.

32. Observational Research

Learning Objectives

1. List the various types of observational research methods and distinguish between each.
2. Describe the strengths and weakness of each observational research method.

What Is Observational Research?

The term **observational research** is used to refer to several different types of non-experimental studies in which behavior is systematically observed and recorded. The goal of observational research is to describe a variable or set of variables. More generally, the goal is to obtain a snapshot of specific characteristics of an individual, group, or setting. As described previously, observational research is non-experimental because nothing is manipulated or controlled, and as such we cannot arrive at causal conclusions using this approach. The data that are collected in observational research studies are often qualitative in nature but they may also be quantitative or both (mixed-methods). There are several different types of observational methods that will be described below.

Naturalistic Observation

Naturalistic observation is an observational method that involves observing people's behavior in the environment in which it typically occurs. Thus naturalistic observation is a type of field research (as opposed to a type of laboratory research). Jane Goodall's famous research on chimpanzees is a classic example of naturalistic observation. Dr. Goodall spent three decades observing chimpanzees in their natural environment in East Africa. She examined such things as chimpanzee's social structure, mating patterns, gender roles, family structure, and care of offspring by observing them in the wild. However, naturalistic observation could more simply involve observing shoppers in a grocery store, children on a school playground, or psychiatric inpatients in their wards. Researchers engaged in naturalistic observation usually make their observations as unobtrusively as possible so that participants are not aware that they are being studied. Such an approach is called **disguised naturalistic observation.** Ethically, this method is considered to be acceptable if the participants remain anonymous and the behavior occurs in a public setting where people would not normally have an expectation of privacy. Grocery shoppers putting items into their shopping carts, for example, are engaged in public behavior that is easily observable by store employees and other shoppers. For this reason, most researchers would consider it ethically acceptable to observe them for a study. On the other hand, one of the arguments against the ethicality of the naturalistic observation of "bathroom behavior" discussed earlier in the book is that people have a reasonable expectation of privacy even in a public restroom and that this expectation was violated.

In cases where it is not ethical or practical to conduct disguised naturalistic observation, researchers can conduct **undisguised naturalistic observation** where the participants are made aware of the researcher presence and monitoring of their behavior. However, one concern with undisguised naturalistic observation is reactivity. **Reactivity** refers to when a measure changes participants' behavior. In the case of undisguised naturalistic observation, the

concern with reactivity is that when people know they are being observed and studied, they may act differently than they normally would. This type of reactivity is known as the **Hawthorne effect**. For instance, you may act much differently in a bar if you know that someone is observing you and recording your behaviors and this would invalidate the study. So disguised observation is less reactive and therefore can have higher validity because people are not aware that their behaviors are being observed and recorded. However, we now know that people often become used to being observed and with time they begin to behave naturally in the researcher's presence. In other words, over time people habituate to being observed. Think about reality shows like Big Brother or Survivor where people are constantly being observed and recorded. While they may be on their best behavior at first, in a fairly short amount of time they are flirting, having sex, wearing next to nothing, screaming at each other, and occasionally behaving in ways that are embarrassing.

Participant Observation

Another approach to data collection in observational research is participant observation. In **participant observation**, researchers become active participants in the group or situation they are studying. Participant observation is very similar to naturalistic observation in that it involves observing people's behavior in the environment in which it typically occurs. As with naturalistic observation, the data that are collected can include interviews (usually unstructured), notes based on their observations and interactions, documents, photographs, and other artifacts. The only difference between naturalistic observation and participant observation is that researchers engaged in participant observation become active members of the group or situations they are studying. The basic rationale for participant observation is that there may be important information that is only accessible to, or can be interpreted only by, someone who is an active participant in the group or situation. Like naturalistic observation, participant observation can be either disguised or undisguised. In **disguised participant observation,** the researchers pretend to be members of the social group they are observing and conceal their true identity as researchers.

In a famous example of disguised participant observation, Leon Festinger and his colleagues infiltrated a doomsday cult known as the Seekers, whose members believed that the apocalypse would occur on December 21, 1954. Interested in studying how members of the group would cope psychologically when the prophecy inevitably failed, they carefully recorded the events and reactions of the cult members in the days before and after the supposed end of the world. Unsurprisingly, the cult members did not give up their belief but instead convinced themselves that it was their faith and efforts that saved the world from destruction. Festinger and his colleagues later published a book about this experience, which they used to illustrate the theory of cognitive dissonance (Festinger, Riecken, & Schachter, 1956)[1].

In contrast with **undisguised participant observation,** the researchers become a part of the group they are studying and they disclose their true identity as researchers to the group under investigation. Once again there are important ethical issues to consider with disguised participant observation. First no informed consent can be obtained and second deception is being used. The researcher is deceiving the participants by intentionally withholding information about their motivations for being a part of the social group they are studying. But sometimes disguised participation is the only way to access a protective group (like a cult). Further, disguised participant observation is less prone to reactivity than undisguised participant observation.

Rosenhan's study (1973)[2] of the experience of people in a psychiatric ward would be considered disguised participant observation because Rosenhan and his pseudopatients were admitted into psychiatric hospitals on the pretense of being

1. Festinger, L., Riecken, H., & Schachter, S. (1956). When prophecy fails: A social and psychological study of a modern group that predicted the destruction of the world. University of Minnesota Press.
2. Rosenhan, D. L. (1973). On being sane in insane places. *Science*, *179*, 250–258.

patients so that they could observe the way that psychiatric patients are treated by staff. The staff and other patients were unaware of their true identities as researchers.

Another example of participant observation comes from a study by sociologist Amy Wilkins on a university-based religious organization that emphasized how happy its members were (Wilkins, 2008)[3]. Wilkins spent 12 months attending and participating in the group's meetings and social events, and she interviewed several group members. In her study, Wilkins identified several ways in which the group "enforced" happiness—for example, by continually talking about happiness, discouraging the expression of negative emotions, and using happiness as a way to distinguish themselves from other groups.

One of the primary benefits of participant observation is that the researchers are in a much better position to understand the viewpoint and experiences of the people they are studying when they are a part of the social group. The primary limitation with this approach is that the mere presence of the observer could affect the behavior of the people being observed. While this is also a concern with naturalistic observation, additional concerns arise when researchers become active members of the social group they are studying because that they may change the social dynamics and/or influence the behavior of the people they are studying. Similarly, if the researcher acts as a participant observer there can be concerns with biases resulting from developing relationships with the participants. Concretely, the researcher may become less objective resulting in more experimenter bias.

Structured Observation

Another observational method is **structured observation**. Here the investigator makes careful observations of one or more specific behaviors in a particular setting that is more structured than the settings used in naturalistic or participant observation. Often the setting in which the observations are made is not the natural setting. Instead, the researcher may observe people in the laboratory environment. Alternatively, the researcher may observe people in a natural setting (like a classroom setting) that they have structured some way, for instance by introducing some specific task participants are to engage in or by introducing a specific social situation or manipulation.

Structured observation is very similar to naturalistic observation and participant observation in that in all three cases researchers are observing naturally occurring behavior; however, the emphasis in structured observation is on gathering quantitative rather than qualitative data. Researchers using this approach are interested in a limited set of behaviors. This allows them to quantify the behaviors they are observing. In other words, structured observation is less global than naturalistic or participant observation because the researcher engaged in structured observations is interested in a small number of specific behaviors. Therefore, rather than recording everything that happens, the researcher only focuses on very specific behaviors of interest.

Researchers Robert Levine and Ara Norenzayan used structured observation to study differences in the "pace of life" across countries (Levine & Norenzayan, 1999)[4]. One of their measures involved observing pedestrians in a large city to see how long it took them to walk 60 feet. They found that people in some countries walked reliably faster than people in other countries. For example, people in Canada and Sweden covered 60 feet in just under 13 seconds on average, while people in Brazil and Romania took close to 17 seconds. When structured observation takes place in the complex and even chaotic "real world," the questions of when, where, and under what conditions the observations will be made,

3. Wilkins, A. (2008). "Happier than Non-Christians": Collective emotions and symbolic boundaries among evangelical Christians. *Social Psychology Quarterly*, 71, 281–301.
4. Levine, R. V., & Norenzayan, A. (1999). The pace of life in 31 countries. *Journal of Cross-Cultural Psychology*, 30, 178–205.

and who exactly will be observed are important to consider. Levine and Norenzayan described their sampling process as follows:

> "Male and female walking speed over a distance of 60 feet was measured in at least two locations in main downtown areas in each city. Measurements were taken during main business hours on clear summer days. All locations were flat, unobstructed, had broad sidewalks, and were sufficiently uncrowded to allow pedestrians to move at potentially maximum speeds. To control for the effects of socializing, only pedestrians walking alone were used. Children, individuals with obvious physical handicaps, and window-shoppers were not timed. Thirty-five men and 35 women were timed in most cities." (p. 186).

Precise specification of the sampling process in this way makes data collection manageable for the observers, and it also provides some control over important extraneous variables. For example, by making their observations on clear summer days in all countries, Levine and Norenzayan controlled for effects of the weather on people's walking speeds. In Levine and Norenzayan's study, measurement was relatively straightforward. They simply measured out a 60-foot distance along a city sidewalk and then used a stopwatch to time participants as they walked over that distance.

As another example, researchers Robert Kraut and Robert Johnston wanted to study bowlers' reactions to their shots, both when they were facing the pins and then when they turned toward their companions (Kraut & Johnston, 1979)[5]. But what "reactions" should they observe? Based on previous research and their own pilot testing, Kraut and Johnston created a list of reactions that included "closed smile," "open smile," "laugh," "neutral face," "look down," "look away," and "face cover" (covering one's face with one's hands). The observers committed this list to memory and then practiced by coding the reactions of bowlers who had been videotaped. During the actual study, the observers spoke into an audio recorder, describing the reactions they observed. Among the most interesting results of this study was that bowlers rarely smiled while they still faced the pins. They were much more likely to smile after they turned toward their companions, suggesting that smiling is not purely an expression of happiness but also a form of social communication.

In yet another example (this one in a laboratory environment), Dov Cohen and his colleagues had observers rate the emotional reactions of participants who had just been deliberately bumped and insulted by a confederate after they dropped off a completed questionnaire at the end of a hallway. The confederate was posing as someone who worked in the same building and who was frustrated by having to close a file drawer twice in order to permit the participants to walk past them (first to drop off the questionnaire at the end of the hallway and once again on their way back to the room where they believed the study they signed up for was taking place). The two observers were positioned at different ends of the hallway so that they could read the participants' body language and hear anything they might say. Interestingly, the researchers hypothesized that participants from the southern United States, which is one of several places in the world that has a "culture of honor," would react with more aggression than participants from the northern United States, a prediction that was in fact supported by the observational data (Cohen, Nisbett, Bowdle, & Schwarz, 1996)[6].

When the observations require a judgment on the part of the observers—as in the studies by Kraut and Johnston and Cohen and his colleagues—a process referred to as **coding** is typically required. Coding generally requires clearly defining a set of target behaviors. The observers then categorize participants individually in terms of which behavior they have engaged in and the number of times they engaged in each behavior. The observers might even record the duration of each behavior. The target behaviors must be defined in such a way that guides different observers to code them in the same way. This difficulty with coding illustrates the issue of interrater reliability, as mentioned in Chapter 4. Researchers are expected to demonstrate the interrater reliability of their coding procedure by having multiple raters

5. Kraut, R. E., & Johnston, R. E. (1979). Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology*, 37, 1539–1553.
6. Cohen, D., Nisbett, R. E., Bowdle, B. F., & Schwarz, N. (1996). Insult, aggression, and the southern culture of honor: An "experimental ethnography." *Journal of Personality and Social Psychology*, 70(5), 945-960.

code the same behaviors independently and then showing that the different observers are in close agreement. Kraut and Johnston, for example, video recorded a subset of their participants' reactions and had two observers independently code them. The two observers showed that they agreed on the reactions that were exhibited 97% of the time, indicating good interrater reliability.

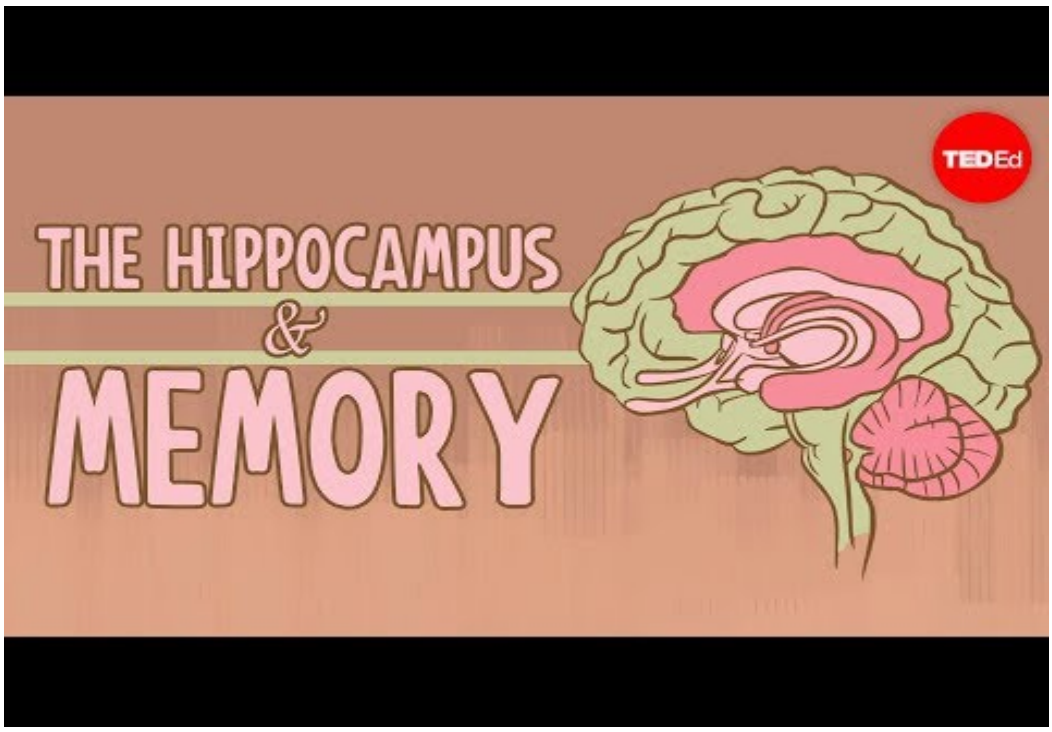
One of the primary benefits of structured observation is that it is far more efficient than naturalistic and participant observation. Since the researchers are focused on specific behaviors this reduces time and expense. Also, often times the environment is structured to encourage the behaviors of interest which again means that researchers do not have to invest as much time in waiting for the behaviors of interest to naturally occur. Finally, researchers using this approach can clearly exert greater control over the environment. However, when researchers exert more control over the environment it may make the environment less natural which decreases external validity. It is less clear for instance whether structured observations made in a laboratory environment will generalize to a real world environment. Furthermore, since researchers engaged in structured observation are often not disguised there may be more concerns with reactivity.

Case Studies

A **case study** is an in-depth examination of an individual. Sometimes case studies are also completed on social units (e.g., a cult) and events (e.g., a natural disaster). Most commonly in psychology, however, case studies provide a detailed description and analysis of an individual. Often the individual has a rare or unusual condition or disorder or has damage to a specific region of the brain.

Like many observational research methods, case studies tend to be more qualitative in nature. Case study methods involve an in-depth, and often a longitudinal examination of an individual. Depending on the focus of the case study, individuals may or may not be observed in their natural setting. If the natural setting is not what is of interest, then the individual may be brought into a therapist's office or a researcher's lab for study. Also, the bulk of the case study report will focus on in-depth descriptions of the person rather than on statistical analyses. With that said some quantitative data may also be included in the write-up of a case study. For instance, an individual's depression score may be compared to normative scores or their score before and after treatment may be compared. As with other qualitative methods, a variety of different methods and tools can be used to collect information on the case. For instance, interviews, naturalistic observation, structured observation, psychological testing (e.g., IQ test), and/or physiological measurements (e.g., brain scans) may be used to collect information on the individual.

HM is one of the most notorious case studies in psychology. HM suffered from intractable and very severe epilepsy. A surgeon localized HM's epilepsy to his medial temporal lobe and in 1953 he removed large sections of his hippocampus in an attempt to stop the seizures. The treatment was a success, in that it resolved his epilepsy and his IQ and personality were unaffected. However, the doctors soon realized that HM exhibited a strange form of amnesia, called anterograde amnesia. HM was able to carry out a conversation and he could remember short strings of letters, digits, and words. Basically, his short term memory was preserved. However, HM could not commit new events to memory. He lost the ability to transfer information from his short-term memory to his long term memory, something memory researchers call consolidation. So while he could carry on a conversation with someone, he would completely forget the conversation after it ended. This was an extremely important case study for memory researchers because it suggested that there's a dissociation between short-term memory and long-term memory, it suggested that these were two different abilities sub-served by different areas of the brain. It also suggested that the temporal lobes are particularly important for consolidating new information (i.e., for transferring information from short-term memory to long-term memory).



Reading in print? Scan this QR code to view the video on your mobile device. Or go to youtu.be/KkaXNvzE4pk

A YouTube element has been excluded from this version of the text. You can view it online here: https://ecampusontario.pressbooks.pub/psychmethods3ecan/?p=101

The history of psychology is filled with influential cases studies, such as Sigmund Freud's description of "Anna O." (see Note 6.1 "The Case of "Anna O."") and John Watson and Rosalie Rayner's description of Little Albert (Watson & Rayner, 1920)[7], who allegedly learned to fear a white rat—along with other furry objects—when the researchers repeatedly made a loud noise every time the rat approached him.

The Case of "Anna O."

Sigmund Freud used the case of a young woman he called "Anna O." to illustrate many principles of his theory of psychoanalysis (Freud, 1961)[8]. (Her real name was Bertha Pappenheim, and she was an early feminist who went on to make important contributions to the field of social work.) Anna had come to Freud's colleague Josef Breuer around 1880 with a variety of odd physical and psychological symptoms. One of them was that for several weeks she was unable to drink any fluids. According to Freud,

7. Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology*, 3, 1–14.
8. Freud, S. (1961). *Five lectures on psycho-analysis*. New York, NY: Norton.

> She would take up the glass of water that she longed for, but as soon as it touched her lips she would push it away like someone suffering from hydrophobia....She lived only on fruit, such as melons, etc., so as to lessen her tormenting thirst. (p. 9)

But according to Freud, a breakthrough came one day while Anna was under hypnosis.

> [S]he grumbled about her English "lady-companion," whom she did not care for, and went on to describe, with every sign of disgust, how she had once gone into this lady's room and how her little dog—horrid creature!—had drunk out of a glass there. The patient had said nothing, as she had wanted to be polite. After giving further energetic expression to the anger she had held back, she asked for something to drink, drank a large quantity of water without any difficulty, and awoke from her hypnosis with the glass at her lips; and thereupon the disturbance vanished, never to return. (p.9)

Freud's interpretation was that Anna had repressed the memory of this incident along with the emotion that it triggered and that this was what had caused her inability to drink. Furthermore, he believed that her recollection of the incident, along with her expression of the emotion she had repressed, caused the symptom to go away.

As an illustration of Freud's theory, the case study of Anna O. is quite effective. As evidence for the theory, however, it is essentially worthless. The description provides no way of knowing whether Anna had really repressed the memory of the dog drinking from the glass, whether this repression had caused her inability to drink, or whether recalling this "trauma" relieved the symptom. It is also unclear from this case study how typical or atypical Anna's experience was.

Figure 10.1 Anna O. "Anna O." was the subject of a famous case study used by Freud to illustrate the principles of psychoanalysis. Source: http://en.wikipedia.org/wiki/File:Pappenheim_1882.jpg

Case studies are useful because they provide a level of detailed analysis not found in many other research methods

and greater insights may be gained from this more detailed analysis. As a result of the case study, the researcher may gain a sharpened understanding of what might become important to look at more extensively in future more controlled research. Case studies are also often the only way to study rare conditions because it may be impossible to find a large enough sample of individuals with the condition to use quantitative methods. Although at first glance a case study of a rare individual might seem to tell us little about ourselves, they often do provide insights into normal behavior. The case of HM provided important insights into the role of the hippocampus in memory consolidation.

However, it is important to note that while case studies can provide *insights* into certain areas and variables to study, and can be useful in helping develop theories, they should never be used as evidence for theories. In other words, case studies can be used as inspiration to formulate theories and hypotheses, but those hypotheses and theories then need to be formally tested using more rigorous quantitative methods. The reason case studies shouldn't be used to provide support for theories is that they suffer from problems with both internal and external validity. Case studies lack the proper controls that true experiments contain. As such, they suffer from problems with internal validity, so they cannot be used to determine causation. For instance, during HM's surgery, the surgeon may have accidentally lesioned another area of HM's brain (a possibility suggested by the dissection of HM's brain following his death) and that lesion may have contributed to his inability to consolidate new information. The fact is, with case studies we cannot rule out these sorts of alternative explanations. So, as with all observational methods, case studies do not permit determination of causation. In addition, because case studies are often of a single individual, and typically an abnormal individual, researchers cannot generalize their conclusions to other individuals. Recall that with most research designs there is a trade-off between internal and external validity. With case studies, however, there are problems with both internal validity and external validity. So there are limits both to the ability to determine causation and to generalize the results. A final limitation of case studies is that ample opportunity exists for the theoretical biases of the researcher to color or bias the case description. Indeed, there have been accusations that the woman who studied HM destroyed a lot of her data that were not published and she has been called into question for destroying contradictory data that didn't support her theory about how memories are consolidated. There is a fascinating New York Times article that describes some of the controversies that ensued after HM's death and analysis of his brain that can be found at: https://www.nytimes.com/2016/08/07/magazine/the-brain-that-couldnt-remember.html?_r=0

Archival Research

Another approach that is often considered observational research involves analyzing archival data that have already been collected for some other purpose. An example is a study by Brett Pelham and his colleagues on "implicit egotism"—the tendency for people to prefer people, places, and things that are similar to themselves (Pelham, Carvallo, & Jones, 2005)[9]. In one study, they examined Social Security records to show that women with the names Virginia, Georgia, Louise, and Florence were especially likely to have moved to the states of Virginia, Georgia, Louisiana, and Florida, respectively.

As with naturalistic observation, measurement can be more or less straightforward when working with archival data. For example, counting the number of people named Virginia who live in various states based on Social Security records is relatively straightforward. But consider a study by Christopher Peterson and his colleagues on the relationship between optimism and health using data that had been collected many years before for a study on adult development (Peterson, Seligman, & Vaillant, 1988)[10]. In the 1940s, healthy male college students had completed an open-ended

9. Pelham, B. W., Carvallo, M., & Jones, J. T. (2005). Implicit egotism. *Current Directions in Psychological Science*, 14, 106–110.
10. Peterson, C., Seligman, M. E. P., & Vaillant, G. E. (1988). Pessimistic explanatory style is a risk factor

questionnaire about difficult wartime experiences. In the late 1980s, Peterson and his colleagues reviewed the men's questionnaire responses to obtain a measure of explanatory style—their habitual ways of explaining bad events that happen to them. More pessimistic people tend to blame themselves and expect long-term negative consequences that affect many aspects of their lives, while more optimistic people tend to blame outside forces and expect limited negative consequences. To obtain a measure of explanatory style for each participant, the researchers used a procedure in which all negative events mentioned in the questionnaire responses, and any causal explanations for them were identified and written on index cards. These were given to a separate group of raters who rated each explanation in terms of three separate dimensions of optimism-pessimism. These ratings were then averaged to produce an explanatory style score for each participant. The researchers then assessed the statistical relationship between the men's explanatory style as undergraduate students and archival measures of their health at approximately 60 years of age. The primary result was that the more optimistic the men were as undergraduate students, the healthier they were as older men. Pearson's *r* was +.25.

This method is an example of **content analysis**—a family of systematic approaches to measurement using complex archival data. Just as structured observation requires specifying the behaviors of interest and then noting them as they occur, content analysis requires specifying keywords, phrases, or ideas and then finding all occurrences of them in the data. These occurrences can then be counted, timed (e.g., the amount of time devoted to entertainment topics on the nightly news show), or analyzed in a variety of other ways.

for physical illness: A thirty-five year longitudinal study. *Journal of Personality and Social Psychology*, 55, 23–27.

33. Key Takeaways and Exercises

Key Takeaways

- Non-experimental research is research that lacks the manipulation of an independent variable.
- There are two broad types of non-experimental research. Correlational research that focuses on statistical relationships between variables that are measured but not manipulated; and observational research in which participants are observed and their behavior is recorded without the researcher interfering or manipulating any variables.
- In general, experimental research is high in internal validity, correlational research is low in internal validity, and quasi-experimental research is in between.
- Correlational research involves measuring two variables and assessing the relationship between them, with no manipulation of an independent variable.
- Correlation does not imply causation. A statistical relationship between two variables, X and Y, does not necessarily mean that X causes Y. It is also possible that Y causes X, or that a third variable, Z, causes both X and Y.
- While correlational research cannot be used to establish causal relationships between variables, correlational research does allow researchers to achieve many other important objectives (establishing reliability and validity, providing converging evidence, describing relationships, and making predictions)
- Correlation coefficients can range from -1 to +1. The sign indicates the direction of the relationship between the variables and the numerical value indicates the strength of the relationship.
- Researchers often use complex correlational research to explore relationships among several variables in the same study.
- Complex correlational research can be used to explore possible causal relationships among variables using techniques such as partial correlation and multiple regression. Such designs can show patterns of relationships that are consistent with some causal interpretations and inconsistent with others, but they cannot unambiguously establish that one variable causes another.
- Qualitative research is an important alternative to quantitative research in psychology. It generally involves asking broader research questions, collecting more detailed data (e.g., interviews), and using non-statistical analyses.
- Many researchers conceptualize quantitative and qualitative research as complementary and advocate combining them. For example, qualitative research can be used to generate hypotheses and quantitative research to test them.
- There are several different approaches to observational research including naturalistic observation, participant observation, structured observation, case studies, and archival research.
- Naturalistic observation is used to observe people in their natural setting; participant observation involves becoming an active member of the group being observed; structured observation involves coding a small number of behaviors in a quantitative manner; case studies are typically used to collect in-depth information on a single individual; and archival research involves analyzing existing data.

Exercises

- Discussion: For each of the following studies, decide which type of research design it is and explain why.
    - A researcher conducts detailed interviews with unmarried teenage fathers to learn about how they feel and what they think about their role as fathers and summarizes their feelings in a written narrative.
    - A researcher measures the impulsivity of a large sample of drivers and looks at the statistical relationship between this variable and the number of traffic tickets the drivers have received.
    - A researcher randomly assigns patients with low back pain either to a treatment involving hypnosis or to a treatment involving exercise. She then measures their level of low back pain after 3 months.
- Discussion: For each of the following, decide whether it is most likely that the study described is experimental or non-experimental and explain why.
    - A cognitive psychologist compares the ability of people to recall words that they were instructed to "read" with their ability to recall words that they were instructed to "imagine."
    - A manager studies the correlation between new employees' college grade point averages and their first-year performance reports.
    - An automotive engineer installs different stick shifts in a new car prototype, each time asking several people to rate how comfortable the stick shift feels.
    - A food scientist studies the relationship between the temperature inside people's refrigerators and the amount of bacteria on their food.
    - A social psychologist tells some research participants that they need to hurry over to the next building to complete a study. She tells others that they can take their time. Then she observes whether they stop to help a research assistant who is pretending to be hurt.
- Practice: For each of the following statistical relationships, decide whether the directionality problem is present and think of at least one plausible third variable.
    - People who eat more lobster tend to live longer.
    - People who exercise more tend to weigh less.
    - College students who drink more alcohol tend to have poorer grades.
- Practice: Construct a correlation matrix for a hypothetical study including the variables of depression, anxiety, self-esteem, and happiness. Include the Pearson's *r* values that you would expect.
- Discussion: Imagine a correlational study that looks at intelligence, the need for cognition, and high school students' performance in a critical thinking course. A multiple regression analysis shows that intelligence is not related to performance in the class but that the need for cognition is. Explain what this study has shown in terms of what is related to good performance in the critical thinking course.
- Discussion: What are some ways in which a qualitative study of girls who play youth baseball would likely differ from a quantitative study on the same topic? How would the data differ by interviewing girls one-on-one rather than conducting focus groups or surveys?
- Practice: Find and read a published case study in psychology. (Use *case study* as a key term in a

PsycINFO search.) Then do the following:

- Describe one problem related to internal validity.
- Describe one problem related to external validity.
- Generate one hypothesis suggested by the case study that might be interesting to test in a subsequent study.