What do you think psychologists do? If you are like most people, when you think of a psychologist, you think of a therapist counseling people who have problems. If I asked you to name a psychologist, you would probably name Sigmund Freud. However, Freud and psychologists who work as therapists are not the focus of this book. They will be discussed, but they are only a part of psychology's story. Psychology is a science, not just a mental health profession. The subjects of this scientific study are you, me, all humans. Some psychologists may use other animals in their research, but their main goal is still predominantly to understand humans. **Psychology** is the science of human behavior and mental processes. Psychologists attempt to understand all aspects of both our observable behavior, such as speech and physical movement, and internal mental processes, such as remembering and thinking, which cannot be directly observed. Psychologists may be found in any number of roles, including teaching, researching, consulting, and yes, counseling troubled people. This book, however, will focus on the research done by psychological scientists, the process by which they've accomplished that research, and what we've learned from their work.

Psychological researchers study everything about us from how our brain works and how we see and hear to how we reason and make decisions. The American Psychological Association lists 54 different divisions of psychology, and psychologists specialize in studying each of these different aspects of our behavior and mental processing. To learn more about these various subfields and careers in psychology, visit www.apa.org/careers/resources/guides/careers.aspx. Although there are many diverse areas within psychology, there are only four major research perspectives for studying these topics. We will begin with a general overview of these four perspectives and then provide descriptions of the major research methods that psychologists use regardless of their perspective. Understanding these perspectives and the research methods used by psychologists will allow you to start thinking like a psychologist (like a scientist).

Note that there are other perspectives in psychology that are primarily clinical in nature (related to psychological therapy). We will discuss the psychoanalytic perspective (which emphasizes the interaction of unconscious forces and childhood experiences in personality development) and the humanistic perspective (which emphasizes the personal growth motive) in Chapter 8, Personality Theories and Assessment, and the psychotherapies based on these two perspectives in Chapter 10, Abnormal Psychology.

The Four Major Research Perspectives

There are four major research perspectives—biological, cognitive, behavioral, and sociocultural. It's important to understand that these perspectives are complementary. The research findings from the major perspectives fit together like the pieces of a jigsaw puzzle to give us a more complete picture. No particular perspective is better than the others, and psychologists using the various perspectives work together to provide a more complete explanation of our behavior and mental processing.

The best way to understand how the major research perspectives differ is to consider the major goal of psychologists—to explain human behavior and mental processes. To explain means to know the causes of our behavior and mental processes. To facilitate your

psychology The science of human behavior and mental processes.

understanding of these perspectives, I discuss them in two different pairs based on the type of causal factors that they emphasize—internal factors or external factors. The biological perspective and the cognitive perspective focus on causes that stem from within us (internal factors); the behavioral perspective and the sociocultural perspective focus on causes that stem from outside us (external factors). We'll also briefly consider developmental psychology, a research field that provides a nice example of how these perspectives complement one another.

Perspectives Emphasizing Internal Factors

The biological perspective and the cognitive perspective focus on internal factors. In the case of the **biological perspective,** our physiological hardware (especially the brain and nervous system) is viewed as the major determiner of behavior and mental processing. The genetic and evolutionary bases of our physiology are also important. In contrast, for the **cognitive perspective,** the major explanatory focus is on how our mental processes, such as perception, memory, and problem solving, work and impact our behavior. To contrast this perspective with the biological perspective, you can think of these mental processes as the software, or programs, of the brain (the wetware, the biological corollary to computer hardware).

The biological perspective. We are biological creatures; therefore, looking for explanations in terms of our biology makes sense. Biological psychologists look for causes within our physiology, our genetics, and human evolution. They argue that our actions and thoughts are functions of our underlying biology. Let's consider an example of what most people would term a "psychological" disorder, depression. Why do we get depressed? A biological psychologist might focus on a deficiency in the activity of certain chemicals in the nervous system as a cause of this problem. Therefore, to treat depression using this perspective, the problem with the chemical deficiency would have to be rectified. How? Antidepressant drugs such as Prozac or Zoloft might be prescribed. These drugs increase the activity of the neural chemicals involved, and this increased activity might lead to changes in our mood. If all goes well, a few weeks after beginning treatment, we begin to feel better. Thus, our mood is at least partly a function of our brain chemistry. Of course, many nonbiological factors can contribute to depression, including unhealthy patterns of thinking, learned helplessness, and disturbing life circumstances. It's important to remember that employing psychology's complementary perspectives in addressing research and clinical issues provides the most complete answer.

In addition to the impact of brain chemistry, biological psychologists also study the involvement of the various parts of the brain and nervous system on our behavior and mental processes. For example, they have learned that our "eyes" are indeed in the back of our head. Biological psychologists have found that it is the back part of our brain that allows us to see the world. So, a more correct expression would be that "our

biological perspective A research perspective whose major explanatory focus is how the brain, nervous system, and other physiological mechanisms produce behavior and mental processes.

cognitive perspective A research perspective whose major explanatory focus is how mental processes, such as perception, memory, and problem solving, work and impact behavior.

eyes are in the back of our brain." The brain is not only essential for vision, but it is also the control center for almost all of our behavior and mental processing. In Chapter 2, Neuroscience, you will learn how the brain manages this incredibly difficult task as well as about the roles of other parts of our nervous system and the many different chemicals that transmit information within it, and in Chapter 3, Sensation and Perception, you will learn how our two major senses, vision and audition, work.

The cognitive perspective. Cognitive psychologists study all aspects of cognitive processing from perception to the higher-level processes, such as problem solving and reasoning. Let's try a brief exercise to gain insight into one aspect of our cognitive processing. I will name a category, and you say aloud as fast as you can the first instance of that category that comes to mind. Are you ready? The first category is FRUIT. If you are like most people, you said apple or orange. Let's try another one. The category is PIECE OF FURNITURE. Again, if you are like most people, you said chair or sofa. Why, in the case of FRUIT, don't people say pomegranate or papaya? How do we have categories of information organized so that certain examples come to mind first for most of us? In brief, cognitive research has shown that we organize categorical information around what we consider the most typical or representative examples of a category (Rosch, 1973). These examples (such as apple and orange for FRUIT) are called prototypes for the category and are retrieved first when we think of the category.

A broader cognitive processing question concerns how memory retrieval in general works. Haven't you been in the situation of not being able to retrieve information from memory that you know you have stored there? This can be especially frustrating in exam situations. Or think about the opposite—an event or person comes to mind seemingly out of the blue. Why? Even more complex questions arise when we consider how we attempt to solve problems, reason, and make decisions. For example, here's a series problem with a rather simple answer, but most people find it very difficult: What is the next character in the series OTTFFSS_? The answer is not "O." Why is this problem so difficult? The progress that cognitive psychologists have made in answering such questions about cognitive processing will be discussed in Chapter 5, on memory, and Chapter 6, on thinking and intelligence (where you can find the answer to the series problem).

Perspectives Emphasizing External Factors

Both the behavioral perspective and the sociocultural perspective focus on external factors in explaining human behavior and mental processing. The **behavioral perspective** emphasizes conditioning of our behavior by environmental events, and there is more emphasis on explaining observable behavior than on unobservable mental processes. The **sociocultural perspective** also emphasizes the influence of the external environment, but it more specifically focuses on the impact of other people and our culture as the major

behavioral perspective A research perspective whose major explanatory focus is how external environmental events condition observable behavior.

sociocultural perspective A research perspective whose major explanatory focus is how other people and the cultural context impact behavior and mental processes.

determiners of our behavior and mental processing. In addition to conditioning, the sociocultural perspective equally stresses cognitive types of learning, such as learning by observation or modeling, and thus focuses just as much on mental processing as observable behavior.

The behavioral perspective. According to the behavioral perspective, we behave as we do because of our past history of conditioning by our environment. There are two major types of conditioning, classical (or Pavlovian) and operant. You may be familiar with the most famous example of classical conditioning—Ivan Pavlov's dogs (Pavlov, 1927/1960). In his research, Pavlov sounded a tone and then put food in a dog's mouth. The pairing of these two environmental events led the dog to salivate to the tone in anticipation of the arrival of the food. The salivary response to the tone was conditioned by the sequencing of the two environmental events (the sounding of the tone and putting food in the dog's mouth). The dog learned that the sound of the tone meant food was on its way. According to behaviorists, such classical conditioning can explain how we learn fear and other emotional responses, taste aversions, and many other behaviors.

Classical conditioning is important in determining our behavior, but behaviorists believe operant conditioning is even more important. Operant conditioning involves the relationship between our behavior and its environmental consequences (whether they are reinforcing or punishing). Simply put, if we are reinforced for a behavior, its probability will increase; if we are punished, the probability will decrease. For example, if you ask your teacher a question and he praises you for asking such a good question and then answers it very carefully, you will tend to ask more questions. However, if the teacher criticizes you for asking a stupid question and doesn't even bother to answer it, you will probably not ask more questions. Environmental events (such as a teacher's response) thus control behavior through their reinforcing or punishing nature. Both types of conditioning, classical and operant, will be discussed in Chapter 4. The point to remember here is that environmental events condition our behavior and are the causes of it.

The sociocultural perspective. This perspective focuses on the impact of other people (individuals and groups) and our cultural surroundings on our behavior and mental processing. We are social animals; therefore other people are important to us and thus greatly affect what we do and how we think. None of us is immune to these social "forces." Haven't your thinking and behavior been impacted by other people, especially those close to you? Our coverage of sociocultural research will emphasize the impact of these social forces on our behavior and mental processing.

To help you understand the nature of sociocultural research, let's consider a famous set of experiments that attempted to explain the social forces operating during a tragic, real-world event—the Kitty Genovese murder in 1964 (Latané & Darley, 1970). Kitty was returning home from work to her apartment in New York City when she was brutally attacked, raped, and stabbed to death with a hunting knife. With the headline "37 Who Saw Murder Didn't Call Police," Martin Gansberg (1964) reported in the *New York Times* that "For more than an hour 38 respectable, law-abiding citizens in Queens watched a killer stalk and stab a woman in three separate attacks in Kew Gardens. . . . Not one person telephoned the police during the assault; one witness called after the woman was dead" (p. 1). A. M. Rosenthal, the *New York Times* editor for Gansberg's article, later developed the story into a short book, *Thirty-Eight Witnesses* (1964). Although recent research (Manning, Levine, & Collins, 2007) has shown that there is no evidence for the key features of this story, such as the large number of witnesses, the story stimulated Bibb Latané and John Darley to explain why the reported 38 witnesses didn't help until it was too late. In fact, according to Levitt and Dubner (2009), the story inspired more research on bystander apathy than on the Holocaust during the 20 years following the murder.

In their attempt to explain why so many people witnessed the murder but didn't help, Latané and Darley manipulated the number of bystanders witnessing the emergencies that they created in their experiments. Their general finding is called the bystander effect (and sometimes the Genovese Syndrome)–the probability of a victim receiving help in an emergency is higher when there is only one bystander than when there are many. In brief, the presence of other bystanders leads us not to help. How did Latané and Darley apply this effect to the Kitty Genovese murder? They explained that social forces present in that situation kept the bystanders from helping. Each felt that someone else would do something and that surely someone else had already called the police. Hence, no one helped. Bystander research, along with studies of other intriguing topics that also involve social forces, such as why we conform and why we obey even when it may lead to destructive behavior, will be detailed in Chapter 9, on social psychology.

Now you have at least a general understanding of the four major research perspectives, summarized in Table 1.1. Remember, these perspectives are complementary, and, when used together, help us to gain a more complete understanding of our behavior and mental processes. Developmental psychology (the scientific study of human development across the lifespan) is a research area that nicely illustrates the benefits of using multiple research perspectives to address experimental questions. A good example is the study of how children acquire language. Initially, behavioral learning principles of reinforcement and imitation were believed to be sufficient to account for language acquisition. Although these principles

Table 1.1 The Four Major Research Perspectives in Psychology

Research Perspective	Major Explanatory Focus
Biological	How our physiology (especially the brain and nervous system) produces our behavior and mental processes and how genetics and evolution have impacted our physiology
Cognitive	How our mental processes, such as perception, memory, and problem solving, work and how they impact our behavior
Behavioral	How external environmental events condition our observable behavior
Sociocultural	How other people and the cultural context impact our behavior and mental processes

clearly do play a role (Whitehurst & Valdez-Menchaca, 1988), most developmental language researchers now recognize that biology, cognition, and the sociocultural context are also critical to language learning (Pinker, 1994; Tomasello, 2003). Studies of the brain, for example, indicate that specific brain areas are involved in language acquisition. Research has also revealed that cognitive processes are important as well. For example, as children acquire new concepts, they learn the names that go with them and thus expand their vocabulary. In addition, it has been shown that the sociocultural context of language helps children to learn about the social pragmatic functions of language. For instance, they use a variety of pragmatic cues (such as an adult's focus of attention) for word learning. Thus, by using all four research perspectives, developmental researchers have gained a much better understanding of how children acquire language. We will learn in Chapter 7, Developmental Psychology, that our understanding of many developmental questions has been broadened by the use of multiple research perspectives.

Subsequent chapters will detail the main concepts, theories, and research findings in the major fields of psychology. As you learn about these theories and research findings, beware of the **hindsight bias (I-knew-it-all-along phenomenon)**—the tendency, after learning about an outcome, to be overconfident in one's ability to have predicted it. Per the old adage, "Hindsight is 20/20." Hindsight bias has been widely studied, having been featured in more than 800 articles (Roese & Vohs, 2012). It has been observed in various countries and among both children and adults (Blank, Musch, & Pohl, 2007). Research has shown that after people learn about an experimental finding, the finding seems obvious and very predictable (Slovic & Fischhoff, 1977). Almost any conceivable psychological research finding may seem like common sense after you learn about it. Hindsight bias can even make a pair of opposite research conclusions both seem obvious (Teigen, 1986). If you were told that research indicates that "opposites attract," you would likely nod in agreement. Isn't it obvious? Then again, if you had been told that research indicates that "birds of a feather flock together," you would probably have also agreed and thought the finding obvious. Be mindful of hindsight bias as you learn about what psychologists have learned about us. It may lead you to think that this information is more obvious and easier than it actually is. You may falsely think that you already know much of the material and then not study sufficiently, leading to disappointment at exam time. The hindsight bias even works on itself. Don't you feel like you already knew about this bias? Incidentally, social psychology researchers have found that birds of a feather DO flock together and that opposites DO NOT attract (Myers, 2013).

Psychologists' conclusions are based upon scientific research and thus provide the best answers to questions about human behavior and mental processing. Whether these answers sometimes seem obvious or sometimes agree with common sense is not important. What is important is understanding how psychologists conduct this scientific research in order to get the best answers to their questions. In the next section, we discuss their research methods.

hindsight bias (I-knew-it-all-along phenomenon) The tendency, after learning about an outcome, to be overconfident in one's ability to have predicted it.

Section Summary

In this section, we learned that there are four major research perspectives in psychology. Two of them, the biological perspective and the cognitive perspective, focus on internal causes of our behavior and mental processing. The biological perspective focuses on causal explanations in terms of our physiology, especially the brain and nervous system. The cognitive perspective focuses on understanding how our mental processes work and how they impact our behavior. The biological perspective focuses on the physiological hardware, while the cognitive perspective focuses more on the mental processes or software of the brain.

The behavioral perspective and the sociocultural perspective emphasize external causes. The behavioral perspective focuses on how our observable behavior is conditioned by external environmental events. The sociocultural perspective emphasizes the impact that other people (social forces) and our culture have on our behavior and mental processing.

Psychologists use all four perspectives to get a more complete explanation of our behavior and mental processing. None of these perspectives is better than the others; they are complementary. Developmental psychology is a research field that nicely illustrates their complementary nature.

We also briefly discussed the hindsight bias, the I-knew-it-all-along phenomenon. This bias leads us to find outcomes as more obvious and predictable than they truly are. You need to beware of this bias when learning the basic research findings and theories discussed in the remainder of this text. It may lead you to think that this information is more obvious and easier to understand and remember than it actually is. It is important that you realize that psychologists have used scientific research methods to conduct their studies, thereby obtaining the best answers possible to their questions about human behavior and mental processing.

ConceptCheck | 1

Explain how the biological and cognitive research perspectives differ in their explanations of human behavior and mental processing.

Explain how the behavioral and sociocultural research perspectives differ in their explanations of human behavior and mental processing.

Research Methods Used by Psychologists

Regardless of their perspective, psychology researchers use the same research methods. These methods fall into three categories–descriptive, correlational, and experimental. The experimental method is used most often because it allows the researcher to explore cause-effect relationships. Remember, the main goal of psychology is to explain (through cause-effect relationships) human behavior and mental processes. However, sometimes researchers can't conduct experiments. For example, it is obviously unethical to set up an experiment testing the effects of passive smoking on children. Who would knowingly subject a group of children to cigarette smoke? In such situations, psychologists can learn a lot by employing the other methods—descriptive and correlational. Researchers can carefully observe and describe the health effects on one child in a family of smokers, or they can study many families in search of relationships (correlations) between parental smoking and childhood infections. These other research methods also provide data for developing hypotheses (testable predictions about cause-effect relationships) to examine in experimental research. We'll discuss the three types of methods in the following order: descriptive, correlational, and experimental.

Descriptive Methods

There are three types of **descriptive methods:** observational techniques, case studies, and survey research. The main purpose of all three methods is to provide objective and detailed descriptions of behavior and mental processes. However, these descriptive data only allow the researcher to speculate about cause-effect relationships—to develop hypotheses about causal relationships. Such hypotheses must then be tested in experiments. With this important limitation in mind, we'll consider the three descriptive methods one at a time.

Observational techniques. Observational techniques exactly reflect their name. The researcher directly observes the behavior of interest. Such observation can be done in the laboratory. For example, children's behavior can be observed using one-way mirrors in the laboratory. However, behavior in the laboratory setting may not be natural. This is why researchers often use **naturalistic observation,** a descriptive research method in which behavior is observed in its natural setting, without the researcher intervening in the behavior being observed. Researchers use naturalistic observation when they are interested in how humans or other animals behave in their natural environments. The researcher attempts to describe both objectively and thoroughly the behaviors that are present and the relationships among these behaviors. There have been many well-known observational studies of other species of animals in their natural habitats. You are probably familiar with some of them—Dian Fossey's study of mountain gorillas in Africa, on which the movie, *Gorillas in the Mist*, was based, and Jane Goodall's study of chimpanzees in Africa (Fossey, 1983; Goodall, 1986). This method is not used only for the observation of other species of animals. Observational studies of human behavior are conducted in many natural settings such as the workplace and school and in social settings such as bars.

Observational techniques do have a major potential problem, though. The observer may influence or change the behavior of those being observed. This is why observers must remain as unobtrusive as possible, so that the results won't be contaminated by their presence. To overcome this possible shortcoming, researchers use participant observation. In **participant observation,** the observer becomes part of the group being observed. Sometimes naturalistic observation studies that start out with unobtrusive observation end up as participant observation studies. For example, Dian Fossey's

descriptive methods Research methods whose main purpose is to provide objective and detailed descriptions of behavior and mental processes.

naturalistic observation A descriptive research method in which the behavior of interest is observed in its natural setting, and the researcher does not intervene in the behavior being observed.

participant observation A descriptive research method in which the observer becomes part of the group being observed.

study of gorillas turned into participant observation when she was finally accepted as a member of the group. However, in most participant observation studies, the observer begins the study as a participant, whether in a laboratory or natural setting. You can think of this type of study as comparable to doing undercover work. A famous example of such a study involved a group of people posing as patients with symptoms of a major mental disorder to see if doctors at psychiatric hospitals could distinguish them from real patients (Rosenhan, 1973). According to Rosenhan, it turned out that the doctors and staff couldn't do so, but the patients could. Once admitted, these "pseudopatients" acted normally and asked to be released. We will find out what happened to them in Chapter 10, on abnormal psychology.

Ralph Hagen/Cartoonstock

Lady studying chimps, chimp studying lady

Case studies. Detailed observation is also involved in a case study. In a **case study,** the researcher studies an individual in depth over an extended period of time. In brief, the researcher attempts to learn as much as possible about the individual being studied. A life history for the individual is developed, and data for a variety of tests are collected. The most common use of case studies is in clinical settings with patients suffering specific deficits or problems. The main goal of a case study is to gather information that will help in the treatment of the patient. The results of a case study cannot be generalized to the entire population. They are specific to the individual that has been studied. However, case study data do allow researchers to develop hypotheses that can then be tested in experimental research. A famous example of such a case study is that of the late Henry Molaison, a person with amnesia (Scoville & Milner, 1957). He was studied by nearly 100 investigators (Corkin, 2002) and is often referred to as the most studied individual in the history of neuroscience (Squire, 2009). For confidentiality purposes while he was alive (he died in 2008 at the age of 82), only his initials, H. M., were used to identify him in the hundreds of studies that he participated in for over five decades. Thus, we will refer to him as H. M. His case will be discussed in more detail in Chapter 5, Memory, but let's consider some of his story here to illustrate the importance of case studies in the development of hypotheses and the subsequent experimental work to test these hypotheses.

For medical reasons, H. M. had his hippocampus (a part of the brain below the cortex) and the surrounding areas surgically removed at a young age. His case study included testing his memory capabilities in depth after the operation. Except for some amnesia for a period preceding his surgery (especially events in the days immediately before the surgery), he appeared to have normal memory for information that he had learned before the operation, but he didn't

case study A descriptive research method in which the researcher studies an individual in depth over an extended period of time.

survey research A descriptive research method in which the researcher uses questionnaires and interviews to collect information about the behavior, beliefs, and attitudes of particular groups of people.

population The entire group of people that a researcher is studying.

sample The subset of a population that actually participates in a research study.

seem to be able to form any new memories. For example, if he didn't know you before his operation, he would never be able to remember your name regardless of how many times you met with him. He could read a magazine over and over again without ever realizing that he had read it before. He couldn't remember what he had eaten for breakfast. Such memory deficits led to the hypothesis that the hippocampus plays an important role in the formation of new memories; later experimental research confirmed this hypothesis (Cohen & Eichenbaum, 1993). We will learn exactly what role the hippocampus plays in memory in Chapter 5. Remember, researchers cannot make cause-effect statements based on the findings of a case study, but they can formulate hypotheses that can be tested in experiments.

Survey research. The last descriptive method is one that you are most likely already familiar with, survey research. You have probably completed surveys either online, over the phone, via the mail, or in person during an interview. **Survey research** uses questionnaires and interviews to collect information about the behavior, beliefs, and attitudes of particular groups of people. It is assumed in survey research that people are willing and able to answer the survey questions accurately. However, the wording, order, and structure of the survey questions may lead the participants to give biased answers (Schwartz, 1999). For example, survey researchers need to be aware of the social desirability bias, our tendency to respond in socially approved ways that may not reflect what we actually think or do. This means that questions need to be constructed carefully to minimize such biases. Developing a well-structured, unbiased set of survey questions is a difficult, time-consuming task, but one that is essential to doing good survey research.

George Price/The New Yorker Collection/The Cartoon Bank

"Next question: I believe that life is a constant striving for balance, requiring frequent tradeoffs between morality and necessity, within a cyclic pattern of joy and sadness, forging a trail of bittersweet memories until one slips, inevitably, into the jaws of death. Agree or disagree?"

Another necessity in survey research is surveying a representative sample of the relevant **population,** the entire group of people being studied. For many reasons (such as time and money), it is impossible to survey every person in the population. This is why the researcher only surveys a **sample,** the subset of people in a population participating in a study. For these sample data to be meaningful, the sample has to be representative of the larger relevant population. If you don't have a representative sample, then generalization of the survey findings to the population is not possible.

One ill-fated survey study of women and love (Hite, 1987) tried to generalize from a nonrepresentative sample (Jackson, 2016). Shere Hite's sample was drawn mainly from women's organizations and political groups, plus some women who requested and completed a survey following the researcher's talk-show appearances. Because such a sample is not representative of American women in general, the results were not either. For example, the estimates of the numbers of women having affairs and disenchanted in their relationships with men were greatly overestimated. To obtain a representative sample, survey researchers usually use random sampling.

In **random sampling,** each individual in the population has an equal opportunity of being in the sample. To understand the "equal opportunity" part of the definition, think about selecting names from a hat, where each name has an equal chance of being selected. In actuality, statisticians have developed procedures for obtaining a random sample that parallel selecting names randomly from a hat. Think about how you would obtain a random sample of first-year students at your college. You couldn't just sample randomly from those first-year students in your psychology class. If you did, all first-year students wouldn't have an equal opportunity to be in your sample. You would have to get a complete list of all first-year students from the registrar and then sample randomly from it. The point to remember is that a survey study must have a representative sample in order to generalize the research findings to the population.

random sampling A sampling technique that obtains a representative sample of a population by ensuring that each individual in a population has an equal opportunity to be in the sample.

correlational study A research study in which two variables are measured to determine if they are related (how well either one predicts the other).

variable Any factor that can take on more than one value.

correlation coefficient A statistic that tells us the type and the strength of the relationship between two variables. The sign of the coefficient (+ or –) indicates the type of correlation—positive or negative, respectively. The absolute value of the coefficient (0.0 to 1.0) represents the strength of the correlation, with 1.0 being the maximum strength.

positive correlation A direct relationship between two variables.

negative correlation An inverse relationship between two variables.

Correlational Studies

In a **correlational study,** two variables are measured to determine if they are related (how well either one predicts the other). A **variable** is any factor that can take on more than one value. For example, age, height, grade point average, and intelligence test scores are all variables. In conducting a correlational study, the researcher first gets a representative sample of the relevant population. Next, the researcher takes the two measurements on the sample. For example, the researcher could measure a person's height and their weight.

The correlation coefficient and predictability. To see if the variables are related, the researcher calculates the **correlation coefficient,** a statistic that tells us the type and the strength of the relationship between two variables. Correlation coefficients range in value from -1.0 to +1.0. The sign of the coefficient, + or -, tells us the type of relationship, positive or negative. A **positive correlation** indicates a direct relationship between two variables—low scores on one variable tend to be paired with low scores on the other variable, and high scores on one variable tend to be paired with high scores on the other variable. Think about the relationship between height and weight. These two variables are positively related. Taller people tend to be heavier. SAT scores and first-year college grades are also positively correlated (Linn, 1982). Students who have higher SAT scores tend to get higher grades during their first year of college.

A **negative correlation** is an inverse relationship between two variables—low scores on one variable tend to be paired with high scores on the other variable, and high scores on one variable tend to be paired with low scores on the other variable. A good example of a negative correlation is the relationship for children between time spent watching television and grades in school—the more time spent watching television, the lower the school grades (Ridley-Johnson, Cooper, & Chance, 1983).

As you know if you have ever climbed a mountain, elevation and temperature are negatively correlated—as elevation increases, temperature decreases. In summary, the sign of the coefficient tells us the type of the relationship between the two variables—positive (+) for a direct relationship or negative (–) for an inverse relationship.

The second part of the correlation coefficient is its absolute value, from 0 to 1.0. The strength of the correlation is indicated by its absolute value. Zero and absolute values near zero indicate no relationship. As the absolute value increases to 1.0, the strength of the relationship increases. Please note that the sign of the coefficient does not tell us anything about the strength of the relationship. Coefficients do not function like numbers on the number line, where positive numbers are greater than negative numbers. With correlation coefficients, only the absolute value of the number tells us about the relationship's strength. For example, –.50 indicates a stronger relationship than +.25. As the strength of the correlation increases, researchers can predict the relationship with more accuracy. If the coefficient is + (or –) 1.0, we have perfect predictability. A correlation of +1.0 means that every change in one variable is accompanied by an equivalent change in the other variable in the same direction, and a correlation of –1.0 means that every change in one variable is accompanied by an equivalent change in the other variable in the opposite direction. Virtually all correlation coefficients in psychological research have an absolute value less than 1.0. Thus, we usually do not have perfect predictability so there will be exceptions to these relationships, even the strong ones. These exceptions do not, however, invalidate the general trends indicated by the correlations. They only indicate that the relationships are not perfect. Because the correlations are not perfect, such exceptions have to exist.

It is also important to realize that when two variables are correlated, but imperfectly so, extreme values on one of the variables tend to be matched on average with less extreme values on the other variable. This is called **regression toward the mean** (the **mean** is the statistical name for the numerical average of a set of values). Probability theory tells us that an extreme value, either far above or below average, is likely to be followed by a value that is more consistent with the long-term average. Regression toward the mean holds for both high and low extreme values. Hence, an extreme low value will move up toward the mean, and an extreme high value will move down toward the mean. A good example of regression toward the mean that you can relate to involves students' exam scores on two exams in a class. The students' scores on the two exams are not perfectly correlated. Regression toward the mean tells us that an unusually high or low score on the first exam will likely be followed by a less extreme score on the second exam. Thus, a student who aced the first exam with a score of 100 will likely get a lower score on the second exam, and a student who failed miserably on the first exam will likely get a higher score on the second exam. There are many factors that impact exam scores for you and other students and thus create this imperfect relationship between the two sets of exam

regression toward the mean The tendency for extreme or unusual values on one variable to be matched on average with less extreme values on the other variable when the two variables are not perfectly correlated.

mean The numerical average for a set of values.

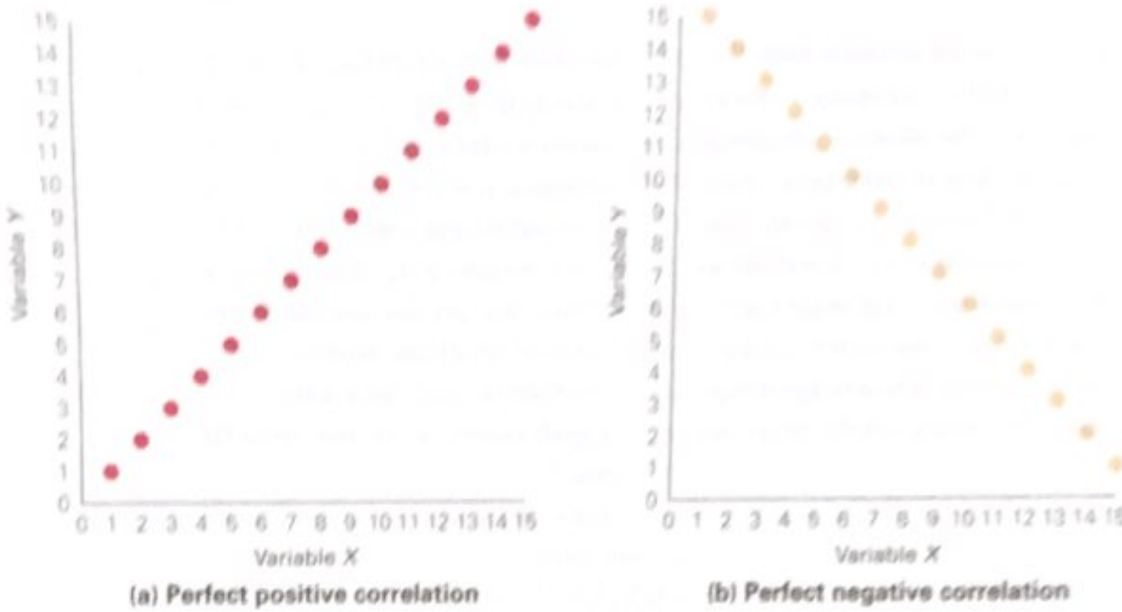
scores. I'm sure that you are familiar with these factors, such as not feeling well when you took one of the exams, studying more for one of the exams, the exam questions overlapping more with the material you emphasized in your study on one exam than the other, and so on. It is important to note that regression toward the mean does not mean that all of the extreme scores have to regress toward the mean but rather that on average the extreme scores will do so. Thus, for example, a student who got a score of 100 on the first exam might also get a perfect score on the second exam. However, if you averaged the scores on the second exam of all of the students who got 100 on the first exam, the average would have regressed toward the mean.

Although regression toward the mean seems straightforward, we do not seem to recognize it when it is operating in our everyday lives. A good example is the widespread belief in the "*Sports Illustrated* jinx"—that it is bad luck for an athlete to be pictured on the cover of *Sports Illustrated* magazine because it will lead to a decline in the athlete's performance afterward (Gilovich, 1991). Many athletes do suffer declines in their performance after being pictured on the cover, but the cause of this decline is regression toward the mean and not a jinx. The athletes pictured on the cover have accomplished outstanding performances in their respective sports. This is why they are on the cover. Regression toward the mean would predict that this extremely high level of performance would be followed by a more normal level of performance because performance levels of an athlete over time vary and thus are not perfectly correlated. Hence, such regression in performance would be expected. Failure to recognize regression toward the mean is not only the reason for belief in the *Sports Illustrated* jinx but also the reason for other superstitious beliefs that people hold and many other events that we all experience in our everyday lives (Kruger, Savitsky, & Gilovich, 1999; Lawson, 2002). Extreme events tend to be followed by more ordinary ones. When this happens, the important point to remember is that regression toward the mean is likely operating, and no extraordinary explanation is necessary. Not only do people fail to recognize regression toward the mean when it occurs, they also tend to make nonregressive or insufficiently regressive predictions (Gilovich, 1991; Kahneman & Tversky, 1973). Thus, when trying to explain or predict changes in scores, levels of performance, and other events over time, remember regression toward the mean is likely operating so extreme or extraordinary events will tend to be followed by more normal ones.

Scatterplots. A good way to understand the predictability of a coefficient is to examine a **scatterplot**—a visual depiction of correlational data. In a scatterplot, each data point represents the scores on two variables for each participant. Several sample scatterplots are presented in Figure 1.1 (page 14). Correlational studies usually involve a large number of participants; therefore, there are usually a large number of data points in a scatterplot. Because those in Figure 1.1 are just examples to illustrate how to interpret scatterplots, there are only 15 points in each one. This means there were 15 participants in each of the hypothetical correlational studies leading to these scatterplots.

The scatterplots in Figure 1.1(a) and (b) indicate perfect 1.0 correlations—(a) a perfect positive correlation and (b) a

scatterplot A visual depiction of correlational data in which each data point represents the scores on the two variables for each participant.



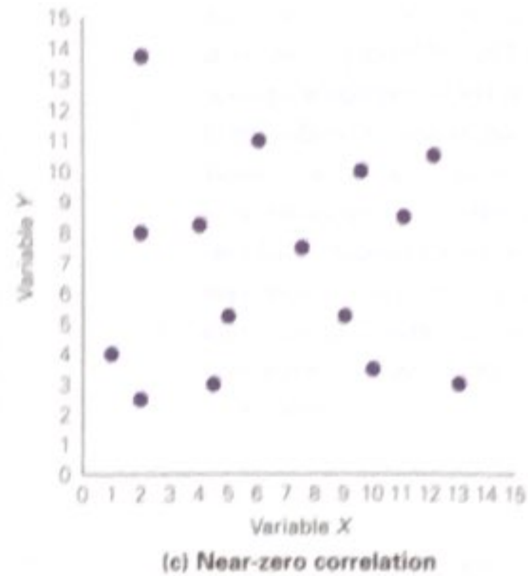
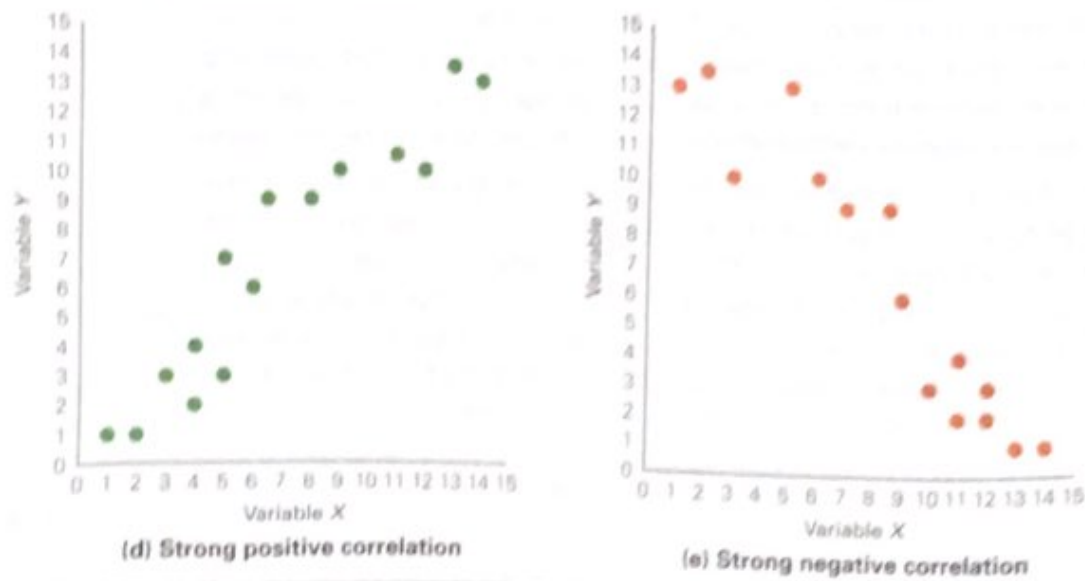


Figure 1.1 | Some Sample Scatterplots | (a) and (b) are examples of perfect correlations because there is no scatter—all of the data points in each plot fall on the same line. The correlation in (a) is positive because the data points show an increasing trend (go from bottom left to top right) and is negative in (b) because the data points show a decreasing trend (go from top left to bottom right). (c) is an example of a near-zero correlation because the data points are scattered all over and do not show a directional trend. (d) is an example of a strong positive correlation because there is not much scatter and the data points have an increasing trend. (e) is an example of a strong negative correlation because there is not much scatter and the data points show a decreasing trend.



perfect negative correlation. All of the points fall on the same line in each scatterplot, which allows us to predict one variable from the other perfectly by using the equation for the line. This means that you have maximal predictability. Please note that the difference between (a) and (b) is the direction of the data points (line). If the data points show an increasing trend (go from the bottom left to the top right of the scatterplot) as in (a), it is a positive relationship. Low scores on one variable tend to be paired with low scores on the other variable, and high scores with high scores. This is a direct relationship. However, if the data points show a decreasing trend (go from the top left to the bottom right) as in (b), there is a negative relationship. Low scores tend to be paired with high scores, and high scores with low scores. This is an inverse relationship.

The scatterplot in Figure 1.1(c) indicates no relationship between the two variables. There is no direction to the data points in this scatterplot. They are scattered all over in a random fashion. This means that we have a correlation near 0 and minimal predictability. Now consider Figure 1.1(d) and (e). First, you should realize that (d) indicates a positive correlation because of the direction of the data points from the bottom left to the top right and that (e) indicates a negative correlation because of the direction of the scatter from the top left to the bottom right. But what else does the scatter of the data points tell us? Note that the data points in (d) and (e) neither fall on the same line as in (a) and (b), nor are they scattered all about the graph with no directional component as in (c). Thus, scatterplots (d) and (e) indicate correlations with strengths somewhere between 0 and 1.0. As the amount of scatter of the data points increases, the strength of the correlation decreases. So, how strong would the correlations represented in (d) and (e) be? They would be fairly strong because there is not much scatter. Remember, as the amount of scatter increases, the predictability also decreases. When the scatter is maximal as in (c), the strength is near 0, and we have little predictability.

The third-variable problem. Strong correlations give us excellent predictability, but they do not allow us to draw cause–effect conclusions about the relationships between the variables. I cannot stress this point enough. Correlation is necessary but not sufficient for causation to exist. Remember, correlational data do not allow us to conclude anything about cause–effect relationships. Only data collected in well-controlled experiments allow us to draw such conclusions. This does not mean that two correlated variables cannot be causally related, but rather that we cannot determine this from correlational data. Maybe they are; maybe they are not. Correlational data do not allow you to decide if they are or are not. Let's see why.

To understand this point, let's consider the negative correlation between self-esteem and depression. As self-esteem decreases, depression increases. But we cannot conclude that low self-esteem causes depression. First, it could be the reverse causal relationship. Isn't it just as likely that depression causes low self-esteem? Second, and of more consequence, isn't it possible that some third factor is responsible for the relationship between the two variables? For example, isn't it possible that some people have a biological predisposition for both low self-esteem and depression or that both self-esteem and depression are the result of a brain chemistry problem?

Both self-esteem and depression could also stem from some current very stressful events. Such alternative possibilities are examples of the **third-variable problem**—another variable may be responsible for the relationship observed between two variables. In brief, such "third variables" are not controlled in a correlational study, making it impossible to determine the cause for the observed relationship.

To make sure you understand the third-variable problem, here is a very memorable example (Li, 1975, described in Stanovich, 2004). Because of overpopulation problems, a correlational study was conducted in Taiwan to identify variables that best predicted the use of contraceptive devices. Correlational data were collected on many different variables, but the researchers found that use of contraceptive devices was most strongly correlated with the number of electrical appliances in the home! Obviously having electrical appliances such as television sets, microwave ovens, and toasters around does not cause people to use birth control. What third variable might be responsible for this relationship? Think about it. A likely one is level of education. People with a higher education tend both to be better informed about birth control and to have a higher socioeconomic status. The former leads them to use contraceptive devices, and the latter allows them to buy more electrical appliances. Charles Wheelan (2013) describes a similar example of the third-variable problem in his book *Naked Statistics.* Consider the positive correlation between a student's SAT scores and the number of television sets that his family owns. Obviously this does not mean that parents can boost their children's test scores by buying another half-dozen television sets. Nor does it likely mean that watching lots of television is good for academic achievement (we have already learned that it isn't). The most logical explanation would be that highly educated people can afford a lot of television sets and tend to have children who test better than average. Thus, both the number of television sets and the test scores are likely the result of a third variable, parental education.

Now that you are aware of the third-variable problem, think about the following finding of a positive correlation between ice cream sales and forest fires (Silver, 2012). What is the third variable driving this correlation? It is the summer heat. Both occur more often in the summer. Correlations that are the results of third variables are called spurious correlations. A **spurious correlation** is one in which the variables are related through their relationship with one or more other

third-variable problem An explanation of a correlation between two variables in terms of another (third) variable that could possibly be responsible for the observed relationship between the two variables.

spurious correlation A correlation in which the variables are related through their relationship with one or more other variables but not through a causal mechanism.

variables but not through a causal mechanism. *Spurious* comes from the Latin word *spurius*, which means false or illegitimate. Let's think through an example of a spurious correlation. There is a positive correlation between shoe size and reading performance in elementary school children. Certainly having bigger feet doesn't cause a child to read better. What's the third variable operating in this example? If you said age, you are right. Older children on average have larger shoe sizes and read better. Tyler Vigen provides hundreds of examples of spurious correlations at his website, www.TylerVigen.com/spurious-correlations and in his book *Spurious Correlations* (2015). Reviewing some of these examples will help to cement your understanding that correlation does not equal causation.

Given the examples of spurious correlations that we have discussed, such as ice cream sales and forest fires, you may think that such correlations do not have any significance in the real world, but they do. To help you understand why they do, I'll describe a famous example from medical research in which the researchers were led astray by a spurious correlation in their search for the cause of a deadly disease, pellagra. In the early 1900s, pellagra victims (pellagrins) had a mortality rate of 1 in 3, and overall, the disease was killing roughly 100,000 people a year in the southeastern United States (Bronfenbrenner & Mahoney, 1975; Kraut, 2003). Medical researchers were scrambling to find its cause so that a treatment could be found. One correlational study found that the incidence of the disease was negatively correlated with the quality of sanitary conditions (Stanovich, 2010). In geographical areas with higher quality plumbing and sewerage, there was hardly any disease, and in areas with poor plumbing and sewerage, there was an extremely high incidence of the disease. Because this correlation fit the infectious disease model in which the disease is spread by a microorganism, these researchers forgot that there could be a spurious correlation stemming from the third-variable problem and concluded that pellagra was being spread by a microorganism in the feces of victims due to poor sanitary conditions. Enter Joseph Goldberger, assigned by the U.S. Surgeon General to determine the cause of pellagra. He realized that the correlation that had been observed could be a spurious one. If so, he asked, what was the third variable? Goldberger's answer—poverty, which was prevalent in much of the southern United States at that time. People living in areas with high-quality sanitary conditions were economically advantaged, and those living in areas with poor-quality sanitary conditions were economically disadvantaged. But how else could poverty lead to the disease, he wondered? Goldberger's answer was that it led to a dietary deficiency that caused the disease. Economically disadvantaged people were eating a corn-based, high-carbohydrate diet with no animal proteins. It consisted mainly of corn, grits, and cornmeal mush with no meat, eggs, or milk, whereas the diet of the economically advantaged was more balanced and included animal proteins.

But why should a dietary deficiency hypothesis be preferred over an infectious disease hypothesis? Both were congruent with the available correlational data. Most medical researchers ignored Goldberger's proposal and continued to pursue the infectious disease hypothesis. Given the urgency of the situation, with thousands dying from the disease, a frustrated Goldberger felt compelled to squash that hypothesis once and for all by exposing himself and some volunteers to its proposed cause in a series of dramatic demonstrations. He, his wife, and 14 volunteers engaged in a series of what were termed "filth parties" because the participants were "feasting on filth" (Kraut, 2003, p. 147). Urine, fecal matter, and liquid feces were taken from pellagrins and then were mixed with wheat flour and formed into dough balls that were swallowed by Goldberger and the volunteers (Kraut, 2003). For complete refutation of the infectious disease hypothesis, they also mixed scabs taken from the rashes of pellagrins into the dough balls, allowed themselves to be injected with blood taken from pellagrins, and applied secretions from the throats and noses of pellagrins with swabs to their own noses and throats. What happened? No one developed pellagra. The infectious disease hypothesis was put to rest. You are likely wondering why Goldberger and these volunteers had so much faith in the dietary deficiency hypothesis that they would participate in these "filthy" demonstrations. Goldberger had collected enough supporting data from dietary studies that he had conducted at orphanages, a state prison, and a state asylum to convince his wife and the volunteers, but not his opponents, that he was right (Kraut, 2003). Sadly, Goldberger died from kidney cancer in 1929 before the specific dietary cause of pellagra, a niacin deficiency, was identified in the 1930s.

In sum, it is important to realize that correlational studies only allow researchers to develop hypotheses about cause-effect relationships. To test these hypotheses so that cause-effect relationships can be determined, researchers conduct experiments in which they manipulate one variable and measure its effect upon another variable while controlling other potentially relevant (third) variables. For example, one of Goldberger's studies preceding the "filth parties" was an experiment with volunteer inmates in a state prison (Kraut, 2003). Goldberger manipulated the diet of the prisoners by having a group of volunteer prisoners eat the corn-based, no-animal-protein diet that he hypothesized caused the disease and measured the incidence of pellagra development in these prisoners. He found that the majority of the 11 prisoners who ate this diet developed pellagra whereas none of the prison population who ate the normal, more balanced prison diet did. Because the only difference between the prisoners was their diets, Goldberger concluded that a dietary deficiency was the cause of pellagra. What happened to the prisoners who developed pellagra? They had volunteered in exchange for pardons after the study ended. Goldberger offered all of those who had developed pellagra treatment until they were well, and most stayed and were successfully treated. However, some refused and left the prison without treatment. Obviously this was an unusual, ethically questionable experiment that involved a serious disease, but it does illustrate the key elements of experimental research—control, manipulation, and measurement. These elements and the experimental research method are the topics we discuss next.

Experimental Research

The key aspect of experimental research is that the researcher controls the experimental setting. The only factor that varies is what the researcher manipulates. It is this control that allows the researcher to make cause-effect statements about the experimental results. This control is derived primarily from two actions. First, the experimenter controls for the possible influence of third variables by making sure that they are held constant across all of the groups or conditions in the experiment. Second, the experimenter controls for any possible influences due to the individual characteristics of the participants, such as intelligence, motivation, and memory, by using **random assignment**—randomly assigning the participants to groups in an experiment in order to equalize participant characteristics across the various groups in the experiment. If the participant characteristics of the groups are on average equivalent at the beginning of the experiment, then any differences between the groups at the end of the experiment cannot be attributed to such characteristics.

random assignment A control measure in which participants are randomly assigned to groups in order to equalize participant characteristics across the various groups in an experiment.

independent variable In an experiment, the variable that is a hypothesized cause and thus is manipulated by the experimenter.

dependent variable In an experiment, a variable that is hypothesized to be affected by the independent variable and thus is measured by the experimenter.

Please note the differences between random assignment and random sampling. Random sampling is a technique in which a sample of participants that is representative of a population is obtained. Hence it is used not only in experiments but also in other research methods such as correlational studies and surveys. Random assignment is only used in experiments. It is a control measure in which the researcher randomly assigns the participants in the sample to the various groups or conditions in an experiment. Random sampling allows you to generalize your results to the relevant population; random assignment controls for possible influences of individual characteristics of the participants on the behavior(s) of interest in an experiment. These differences between random sampling and random assignment are summarized in Table 1.2.

Designing an experiment. When a researcher designs an experiment, the researcher begins with a hypothesis (the prediction to be tested) about the cause-effect relationship between two variables. One of the two variables is assumed to be the cause, and the other variable is the one to be affected. The **independent variable** is the hypothesized cause, and the experimenter manipulates it. The **dependent variable**

Table 1.2 Differences Between Random Sampling and Random Assignment

Random Sampling	Random Assignment
A sampling technique in which a sample of participants that is representative of the population is obtained	A control measure in which participants in a sample are randomly assigned to the groups or conditions in an experiment
Used in experiments and some other research methods such as correlational studies and surveys	Used only in experiments
Allows researcher to generalize the findings to the relevant population	Allows researcher to control for possible influences of individual characteristics of the participants on the behavior(s) of interest

is the variable that is hypothesized to be affected by the independent variable and thus is measured by the experimenter. Thus, in an **experiment** the researcher manipulates the independent variable and measures its effect on the dependent variable while controlling other potentially relevant variables. If there is a causal relationship between the independent and dependent variables, then the measurements of the dependent variable are dependent on the values of the independent variable, hence the name dependent variable. Sometimes the researcher hypothesizes more than one cause or more than one effect so he manipulates more than one independent variable or measures more than one dependent variable. To help you understand this terminology and the mechanics of an experiment, I'll describe an example.

Let's consider the simplest experiment first—only two groups. For control purposes, participants are randomly assigned to these two groups. One of the groups will be exposed to the independent variable, and the other will not. The group exposed to the independent variable is called the **experimental group,** and the group not exposed to the independent variable is called the **control group.** Let's say the experimenter's hypothesis is that aerobic exercise reduces anxiety. The independent variable that will be manipulated is aerobic exercise, and the dependent variable that will be measured is level of anxiety. The experimental group will participate in some aerobic exercise program, and the control group will not. To measure any possible effects of the aerobic exercise on anxiety, the experimenter must measure the anxiety levels of the participants in each group at the beginning of the study before the independent variable is manipulated, and then again after the manipulation. If the two groups are truly equivalent, the level of anxiety for each group at the beginning of the study should be essentially the same. If the aerobic exercise does reduce anxiety, then we should see this difference in the second measurement of anxiety at the end of the experiment.

The independent and dependent variables in an experiment must be operationally defined. An **operational definition** is a description of the operations or procedures the researcher uses to manipulate or measure a variable. In our sample experiment, the operational definition of aerobic exercise would include the type and the duration of the activity. For level of anxiety, the operational definition would describe the way the anxiety variable was measured (for example, a participant's score on a specified anxiety scale). Operational definitions not only clarify a particular experimenter's definitions of variables but also allow other experimenters to attempt to replicate the experiment more easily, and replication is the cornerstone of science (Moonesinghe, Khoury, & Janssens, 2007).

Let's go back to our aerobic exercise experiment. We have our experimental group and our control group, but this experiment really requires a second control group. The first control group (the group not participating in the aerobic exercise program) provides a baseline level of anxiety to which the anxiety of the experimental group can then be compared. In other words, it controls for changes over time in the level of anxiety not due to

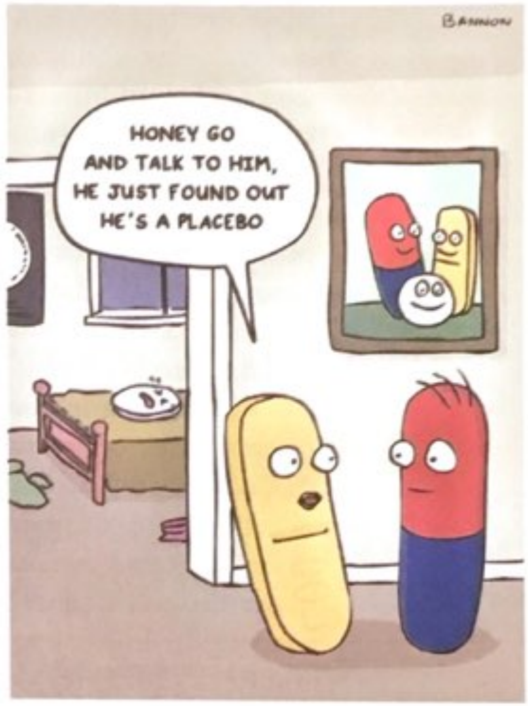
experiment A research method in which the researcher manipulates one or more independent variables and measures their effect on one or more dependent variables while controlling other potentially relevant variables.

experimental group In an experiment, the group exposed to the independent variable.

control group In an experiment, the group not exposed to the independent variable.

operational definition A description of the operations or procedures that a researcher uses to manipulate or measure a variable.

aerobic exercise, such as regression toward the mean. However, we also need to control for what is called the **placebo effect**–improvement due to the expectation of improving because of receiving treatment. The treatment involved in the placebo effect, however, only involves receiving a **placebo**–an inactive pill or a sham treatment that has no known effects. The placebo effect can arise not only from a conscious belief in the treatment but also from subconscious associations between recovery and the experience of being treated (Niemi, 2009). For example, stimuli that a patient links with getting better, such as a doctor's white lab coat or the smell of an examining room, may induce some improvement in the patient's condition. In addition, giving a placebo a popular medication brand name, prescribing more frequent doses, using placebo injections rather than pills, or indicating that it is expensive can boost the effect of a placebo (Niemi, 2009; Stewart-Williams, 2004; Waber, Shiv, Carmon, & Ariely, 2008). Recent research has even found that the placebo effect may not require concealment or deception (Kaptchuk et al., 2010); in the case of pain, that its size may be impacted by personality traits, such as altruism and resilience (Peciña et al., 2013); and for migraine headache drugs, that it may account for greater than 50% of the drug's effectiveness (Kam-Hansen et al., 2014). Given such findings, it is not surprising that almost half of the physicians recently surveyed reported that they prescribed placebo treatments on a regular basis, and over 60% of the surveyed physicians found this to be an ethically permissible practice (Tilburt, Emanuel, Kaptchuk, Curlin, & Miller, 2008).



You might be puzzled by physicians prescribing placebos, but consider the following. Combining the research finding that the placebo response to pain is very large (Kam-Hansen et al., 2014) with the fact that painkiller addiction has become a major problem in our society, suggests that prescribing placebos might be a viable first course of action for the treatment of pain (Marchant, 2016). Given this large placebo effect, a prescription for an active painkiller might then not be necessary, lessening the possibility of addiction. A recent finding that placebo effects for chronic pain have gotten stronger over time in the United States though not elsewhere in the world (Tuttle et al., 2015) makes the use of placebos in the treatment of pain in this country even more viable. According to Tuttle et al., clinical trials in 2013 for drugs to treat chronic pain produced on average only 9% more pain relief than placebos. In 1996, it was 27%. You may be wondering why the placebo effect for chronic pain has become so strong. One likely explanatory factor is that we have come to have a stronger belief in the effectiveness of painkilling drugs because of the media advertising the effectiveness of these drugs and drugs in general, thereby leading us to have a stronger placebo effect because belief plays a crucial role in the

placebo effect Improvement due to the expectation of improving because of receiving treatment.

placebo An inactive pill or a sham treatment that has no known effects.

placebo effect for pain (Benedetti, 2014). It is difficult, for example, to watch commercial television without being bombarded with such drug advertisements. But don't other countries have such advertising? No, such direct-to-consumer advertising for prescription drugs is only legal here in the United States and in New Zealand. There is also evidence for increased placebo responses in clinical trials for some other drugs, such as antidepressants (Rief et al., 2009). Consequently, prescribing placebos as the first step in treatment of medical problems, such as chronic pain, and mental disorders, such as depression, instead of drugs that are addictive or have serious side effects seems to be a very promising possibility.

The placebo (Latin for "I will please) effect, however, does have an evil twin, the **nocebo** (Latin for "I will harm") **effect**, whereby expectation of a negative outcome due to treatment leads to adverse effects (Benedetti, Lanotte, Lopiano, & Colloca, 2007; Kennedy, 1961). Expectations can also do harm; hence, the nocebo effect is sometimes referred to as a negative placebo effect. For example, when a patient anticipates a drug's possible side effects, he can suffer them even if the drug is a placebo. Because of ethical reasons, the nocebo effect has not been studied nearly as much as the placebo effect, but recently Häuser, Hansen, and Enck (2012) reviewed 31 studies that clearly demonstrated the nocebo effect in clinical practice. They concluded that the nocebo effect creates an ethical dilemma for physicians and therapists—how do they fully inform patients of the potential complications of treatment and at the same time minimize inducing these complications through nocebo effects? Similarly, Petersen et al. (2014) reviewed 10 studies on the nocebo effect on pain and found the effect sizes to be similar to those for placebo effects (moderate to large), thereby indicating the importance of minimizing these effects in actual clinical practice. Thus, research on the nocebo effect has mainly centered on its critical role in clinical practice, whereas much of the research on the placebo effect has been concerned with controlling for improvement due to participant expectations in clinical drug trials and experimental studies, as in our aerobic exercise experiment.

The reduction of anxiety in the experimental group participants in our aerobic exercise experiment may be partially or completely due to a placebo effect. This is why researchers add a control group called the placebo group to control for the possible placebo effect. A **placebo group** is a group of participants who believe they are receiving treatment, but they are not. They get a placebo. For example, the participants in a placebo group in the aerobic exercise experiment would be told that they are getting an antianxiety drug, but they would only get a placebo (in this case, a pill that has no active ingredient). The complete design for the aerobic exercise experiment, including the experimental, placebo, and control groups, is shown in Figure 1.2 For the experimenter to conclude that there is a placebo effect, the reduction of anxiety in the placebo group would have to be significantly greater than the reduction for the control group. This is because the control group provides a baseline of the reduction in symptom severity over time caused by factors other than that caused

nocebo effect A negative placebo effect due to the expectation of adverse consequences from receiving treatment.

placebo group A control group of participants who believe they are receiving treatment, but who are only receiving a placebo.

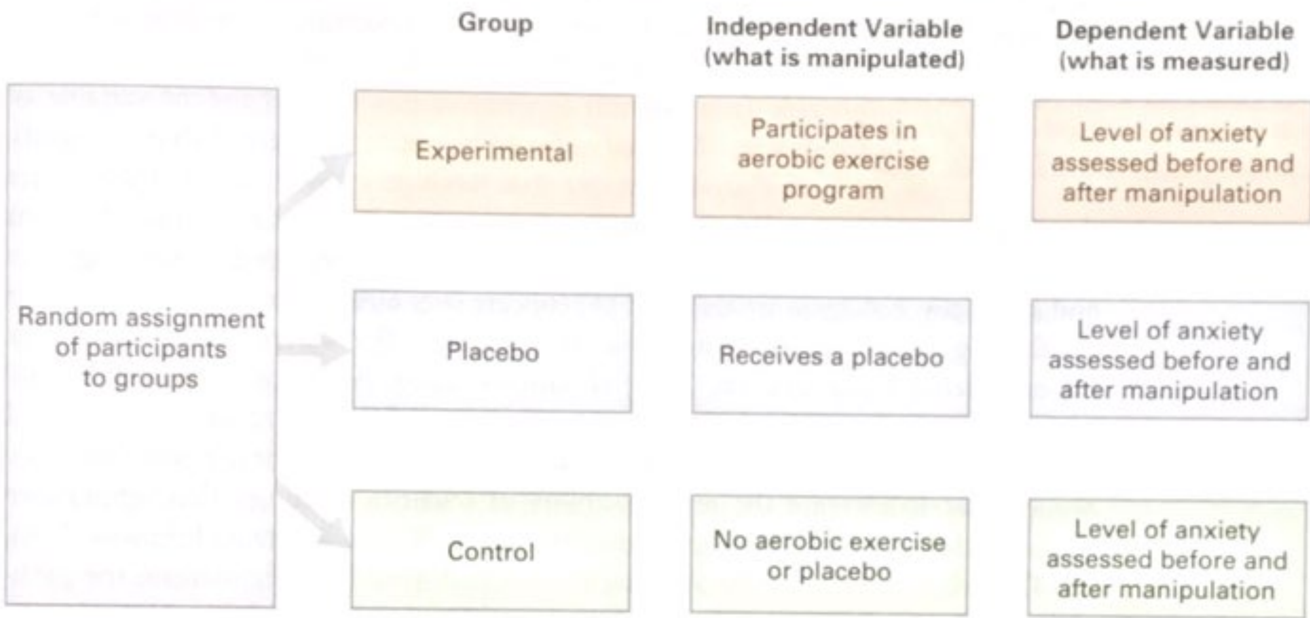


Figure 1.2 | **Design of Aerobic Exercise and Anxiety Experiment** | Participants are randomly assigned to groups in order to equalize participant characteristics across the groups. The placebo group controls for the placebo effect, and the control group provides a baseline level of anxiety reduction for participants who do not participate in the aerobic exercise program or receive a placebo. The level of anxiety reduction for each group is determined by comparing the measurements of the dependent variable (level of anxiety) before and after manipulating the independent variable (aerobic exercise).

by the placebo (Benedetti, 2014). A good example of such a factor is one that we discussed earlier in the chapter, regression toward the mean. In this example, it would translate to your anxiety regressing naturally toward its normal level over time. If a control group was not included in the study, then the measurement of the placebo effect would be confounded because it would include the improvement due to these other factors and not just the placebo. Similarly, for the experimenter to conclude that the reduction of anxiety in the experimental group is due to aerobic exercise and not a placebo effect or other factors leading to a reduction in anxiety, the reduction would have to be significantly greater than that observed for the placebo and control groups because the reduction in the experimental group is due not only to the effect of the aerobic exercise but also to a placebo effect created by the expectation of getting better by engaging in the aerobic exercise and to natural factors leading the body to reduce anxiety and return to its normal state.

Now you may be wondering what is meant by "significantly greater." This is where statistical analysis enters the scene. We use what are called **inferential statistical analyses**—statistical analyses that allow researchers to draw conclusions about the results of their studies. Such analyses tell the researcher the probability that the results of the study are due to random

inferential statistical analyses Statistical analyses that allow researchers to draw conclusions about the results of a study by determining the probability that the results are due to random variation (chance). The results are statistically significant if this probability is .05 or less.

double-blind procedure A control measure in an experiment in which neither the experimenters nor the participants know which participants are in the experimental and control groups.

variation (chance). Obviously, the experimenter would want this probability to be low. Remember, the experimenter's hypothesis is that the manipulation of the independent variable (not chance) is what causes the dependent variable to change. The scientific community has agreed that a "significant" finding is one that has a probability of .05 (1/20) or less that it is due to chance. Possibly, this acceptance standard is too lenient. A recent study that attempted to replicate 100 research findings reported in top cognitive and social psychology journals failed to replicate over 60% of the findings, and when the findings were replicated, the size of the observed effects were on average smaller (Open Science Collaboration, 2015; cf. Gilbert, King, Pettigrew, & Wilson, 2016). Although many factors are involved in this large-scale failure to replicate, a statistical significance standard that may be too lenient is one of them. Some researchers have argued that to increase the reproducibility of scientific findings, this significance level should be raised to a more stringent level, .005 or less (e.g., Johnson, 2013, 2014). Such a revision to the acceptance standard would greatly increase the probability that a finding is not due to chance, and hence its replicability. Nevertheless, the acceptance standard for significance at present remains .05 or less.

Statistical significance tells us the probability that a finding did not occur by chance, but it does not insure that the finding has practical significance or value in our everyday world. A statistically significant finding with little practical value sometimes occurs when very large samples are used in a study. With such samples, very small differences among groups may be significant. Belmont and Marolla's (1973) finding of a birth-order effect for intelligence test scores is a good example of such a finding. Belmont and Marolla analyzed intelligence test data for almost 400,000 19-year-old Dutch males. There was a clear birth-order effect: First borns scored significantly higher than second borns, second borns higher than third borns, and so on. However, the score difference between these groups was very small (only a point or two) and thus not of much practical value. So remember, statistically significant findings do not always have practical significance.

Sidney Harris/ScienceCartoonsPlus.com

"It was more of a 'triple-blind' test. The patients didn't know which ones were getting the real drug, the doctors didn't know, and I'm afraid, nobody knew."

The aerobic exercise experiment would also need to include another control measure, the double-blind procedure. In the **double-blind procedure,** neither the experimenters nor the participants know which participants are in the experimental and control groups. This procedure is called "double-blind" because both the experimenters and participants are blind to (do not know) the participant group assignments. It is not unusual for participants to be blind to which group they have been assigned. This is especially critical for